REPORT DOCUMENTATION PAGE

Form Approved OMB No. 0704-0188

1. REPORT DATE *(DD-MM-YYYY)*	2. REPORT TYPE	3. DATES COVERED *(From – To)*
01-05-2006	Final Report	01-Apr-05 - 01-May-06

4. TITLE AND SUBTITLE

Theoretical Investigation of the Feasibility of PTD-Mediated Translocation of Proteins Across Artificial Membranes

5a. CONTRACT NUMBER
STCU Registration No: P211

5b. GRANT NUMBER

5c. PROGRAM ELEMENT NUMBER

6. AUTHOR(S)

Professor Valeriy N Kharkyanen

5d. PROJECT NUMBER

5d. TASK NUMBER

5e. WORK UNIT NUMBER

7. PERFORMING ORGANIZATION NAME(S) AND ADDRESS(ES)

Fund for scientists of molecular cybernetics and informatics
IOP, 46 Nauki Prospekt
Kiev 03022
Ukraine

8. PERFORMING ORGANIZATION REPORT NUMBER

N/A

9. SPONSORING/MONITORING AGENCY NAME(S) AND ADDRESS(ES)

EOARD
PSC 821 BOX 14
FPO 09421-0014

10. SPONSOR/MONITOR'S ACRONYM(S)

11. SPONSOR/MONITOR'S REPORT NUMBER(S)
STCU 04-8003

12. DISTRIBUTION/AVAILABILITY STATEMENT

Approved for public release; distribution is unlimited.

13. SUPPLEMENTARY NOTES

14. ABSTRACT

This report results from a contract tasking Fund for scientists of molecular cybernetics and informatics as follows: The recent discovery of the ability of protein transduction domains (PTDs) and their synthetic analogues to transport high-molecular weight compounds through biological or artificial membranes is very promising for many applications. However, the mechanism of the phenomenon remains highly debatable and unclear even in its basic features. This project is aimed at the theoretical investigation of the general principles and physical mechanisms of membrane translocation of macromolecules. Modern methods of theoretical physics and computer simulation will be used for the construction and analysis of plausible models. The primary model subject will be the well-characterized translocation domain of the HIV TAT protein. Possible uptake/transduction mechanisms in different models (such as inverted micelle, stochastic membrane pore, etc) will be substantiated or eliminated, with experimentally measured, quantitative evaluation of the translocation efficiency.

15. SUBJECT TERMS

EOARD, Biological And Medical Sciences, Biochemistry

16. SECURITY CLASSIFICATION OF:			17. LIMITATION OF ABSTRACT	18, NUMBER OF PAGES	19a. NAME OF RESPONSIBLE PERSON
a. REPORT	b. ABSTRACT	c. THIS PAGE			ROBERT N. KANG, Lt Col, USAF
UNCLAS	UNCLAS	UNCLAS	UL	77	19b. TELEPHONE NUMBER *(Include area code)* +44 (0)20 7514 4437

Standard Form 298 (Rev. 8/98)
Prescribed by ANSI Std. Z39-18

EOARD Partner Project 211 – Deliverable 1

1. Introduction

One of the main tasks of functional genomics is the study of the role/activity of a given expressed/purified protein *in vivo*, for which it should be inserted in cultured cells or an organism. Initially, this problem was solved in an indirect way based on the transfection of cells with a proper active viral DNA (construction of a proper DNA-vector) with subsequent production of the given protein by cell's own. Later, they started to manage without the DNA-transfection stage, trying to deliver the protein into the cell directly. The first protein delivery methods, however, were the same as those for DNA, such as electroporation, microinjection, cationic lipids, and others. These are, in particular, too invasive and low-efficient. Thus, electroporation often entails ruining a considerable number of cells; microinjection (insertion of femtoliter volumes into a certain cell compartment) needs too specific equipment and sophisticated skill while only single cells are processed, so on. In addition, these methods are suitable for laboratory experiments but are practically inapplicable *in vivo* for their potentially high toxicity, immunogenicity, etc.

That is why in the last decade a great enthusiasm was sparked to the discovered ability of some proteins (and then their excised sequences) to be internalized into cells by themselves and, moreover, to deliver there other macromolecular components, often much larger than the carrier. Early reports about such a peptide appeared in late eighties with the example of the protein-transactivator of transcription of the human immunodeficiency virus, HIV-1 Tat protein [1-3]. In those works it was already said about great potentiality of the phenomenon for the intracellular delivery of macromolecules previously thought to be impermeable to cellular membranes; the evidences of that were reported soon, see e.g. work [4] in which Fawell *et al* showed the entry of proteins cross-linked with relatively large (at that time) Tat-peptides into cultured tissue cells. In a sort, a boom was initiated by the reports of Dowdy *et al* in late nineties (see e.g. [5]) about successful and fast delivery even of large (e.g. β-galactozidase, Mw 120 kDa) proteins fused with a small peptide (Tat-protein fragment) into all mouse tissues, brain included, after intraperitoneal injection. In numerous subsequent works one can trace both a considerable extensive progress – the number of such oligopeptides called protein transduction domains (PTDs) has grown from a few (HIV-1 Tat PTD, *Drosophila* Antennapedia homeodomain penetratin, or Antp, and some others) to about a hundred, as is noted in the recent mini-review [6] – and a diversity of serious controversies, presently culminating to be resolved. The reason of such a sharp interest to the problem is obvious at least from its applied aspect. From a number of works it followed that the translocation of such peptides and their cargo seemed practically non-

specific and almost universal (either to the cell type or cargo), did not go in the classical receptor/endocytosis ways and, in addition, proceeded much faster in comparison with conventionally used methods. Also, the degree of risk or toxicity turned out to be incomparably lower. In the popular scientific literature and on the sites of commercial producers they started to speak about a new era in pharmacology and therapy ([7-8], see also [9]). On the other hand, the phenomenon looked absolutely incomprehensible at least from the commonly accepted physicochemical views on the impermeability of lipid membranes for such highly charged/hydrophilic molecules as cell-penetrating peptides (CPPs).[1] Besides, a series of results have been called into question for unreliability/limitations of the observation methods, especially those that denied the endocytic pathways of the translocation; that is, called into question was the very phenomenon of protein transduction.

The goal of the present deliverable is to review the current literature in the field and select those of the data available that can serve as a starting point in the physical analysis and computer modeling of the transduction mechanisms as the latter aspects, despite of the problem importance, still remain practically inelaborate.

2. Cell-penetrating peptides: general features (outline)

Langel *et al* [6,11] relate the term CPP to up to thirty amino acid amphiphilic peptides which can be internalized by cells by energy-independent mechanisms (although the endocytic pathway is not exluded). According to their classification, CPPs can be conventionally divided into three classes: protein derived CPPs, model peptides and designed peptides. The first are in fact PTDs (with Tat and Antp being most studied), the second represent sequences mimicking the structures of known CPPs (typical are oligoarginines, e.g. $(R)_{n=7\div9}$ [12]); lastly, the third comprise chimeras composed of hydrophobic and hydrophilic parts of different origin (e.g. transportan consisting of a part of galanin attached to mastoparan via lysine).

The primary structure of known CPPS exhibits no special common features apart from the fact that all of them are positively charged and amphipathic (except polycationic homopolymers). Their net positive charge is ensured mainly by arginines and, to lower extent, lysines; thus, (+8)-charged basic domain $Tat_{49\text{-}57}$ RKKRRQRRR contains six arginine and two lysine residues.

There are controversial data on the role of the ***secondary structure***. Thus, Derossi *et al* [13] revealed that internalization of Antp is not hampered by introducing up to three prolines into

[1] Most general term comprising both PTDs and model or chimeric oligopeptides having the mentioned properties [10].

the sequence, i.e. by disruption of the α-helical structure. Together with demonstration of the translocation of reverse helices and D-enantiomers, this allowed the authors to conclude that the internalization mechanism is receptor-independent (but see [14]). For model CPPs studied in [12] (in particular, $(R)_9$ and $(r)_9$) the translocation efficiency is somewhat different, but the main emphasis was put on the guanidine specificity of the arginine structure and the length and conformational flexibility of side chains due to alkyl spacers. The data of Zaro and Shen [15] also show the importance of guanidine structure for transduction while mainly the number of positive charges determines endocytosis. The advantage of arginine-rich peptides over lysine-rich ones and minor importance of other structural characteristics are rather frequently mentioned (see e.g. [16,17]),[2] though there are some reports on better efficiency of lysine homopolymers [18,19]. On the whole, so far it is hard to come to a more or less definite conclusion on the translocation structural preferences.

Even greater diversity of data can be reviewed in regard to the types of cells, unnatural lipid vesicles and cargoes transported by CPPs. Overwhelming majority of experiments on successful translocation is performed *in vitro* on cultured mammalian cells grown by standard methods and then incubated with solutions of CPPs. As for cargoes, their spectrum spreads from small peptides and oligonucleotides (see e.g. [20]) to proteins of 120-150 kDa [5,21] and liposomes of hundreds of nanometers [22] (see also Tables 2,3 in [6]). They can be attached in different ways (usually via a covalent bond); most typical are shown in Fig.1 taken from paper [6].

Binding and translocation mechanisms. As noted above, this central question still remains unclear and gives rise to sharp contradictions. Starting from early works [2,24,25], an intrigue has been generated by reports about bypassing classical endocytic pathways and supposedly direct interaction of CPPs with the lipid bilayer of cellular membranes. The main argument in favor of such statement was the preserved ability (of either TAT PTD or Antp) of translocation at low (4 C) temperature and also under removing ATP sources of energy or in the presence of different inhibitors of endocytosis. Reports of this kind accumulated till 2003 [13,16,17,18,22,26], involving, apart from TAT PTD and Antp, the growing number of different CPPs, although opinions in favor of endocytosis were also expressed (e.g. [27]).

In 2003, however, the situation seemed to be changed due to appearance of several works (first of all [28]) pointed at serious imperfections of the most widespread methods (cell fixation, flow cytometry, etc) resulting in the fact that CPPs, strongly interacting with the cell surface, were not removed even by repeated washings. Consequently, flow cytometry does not

[2] However, the number of arginines, as noted in [16], should not be excessive; thus, for polyarginines $n = 8$ is opimal, with internalization being practically abolished at $n > 12$.

discriminate between internalized and membrane-associated peptides, and this of course leads to overrated indices of penetration and other artefacts and wrong conclusions. The authors of paper [28] insisted that in live, non-fixed cells a typical endosomal distribution of peptides could only be observed and the uptake is inhibited at low temperature and/or ATP depletion. Similar results have been obtained in a series of other works of that year [29-32].

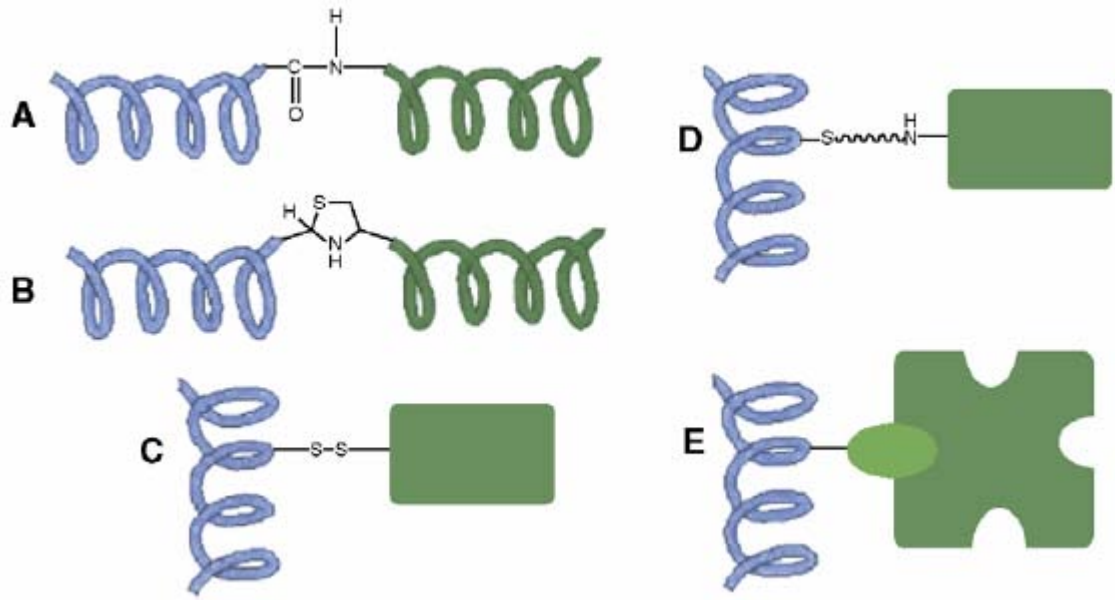


Fig.1. Attachment of cargoes (green) to CPPs (blue). *A,B,C* and *D* represent cargo covalently bound to the CPP via a peptide bond, thiazolidine ring, disulphide bridge and bifunctional linker molecule, respectively. (*E*) A large cargo molecule (e.g. streptavidin, shown in dark green) is non-covalently bound to a smaller cargo (e.g. biotin, light green) that is covalently attached to the CPP [6].

Meanwhile, in the same year at least two papers appeared [15,39] whose authors took the criticism expressed in [28-32] into account and, nevertheless, came to an unambiguous conclusion in favor of the existence of an energy-independent non-endocytic channel of internalization at least for several CPPs (e.g. $(R)_7$), though for TAT PTD and Antp this channel was found to be somewhat restricted. In work [15] with the help of a newly proposed method of subcellular fractionation the contributions of transduction and endocytosis were separated, with the transduction ability being observed for different polyarginine peptides and Tat_{47-57}. As improbable, the endocytic pathway is also noted in works [34,35].

To the moment, one can register a temporal dynamical equilibrium of opinions on this central question. As follows from recent mini-review [6], the transduction ability of CPPs, although remaining incomprehensible so far, is doubtless – the authors of [6] in fact include it into the very definition of CPPs. On the other hand, for many CPPs the role of endocytosis is far from negligible – this is being confirmed by new evidences [36-39], – though the endocytic modes could be non-classical (e.g. caveolar, lipid raft, macropinocytosis, so on [29,40]).

There are relatively less data on more detailed models of CPP membrane binding and translocation. More or less convincingly, one can assert that an important role in the CPP binding to the cellular surface is played by ubiquitous proteoglycans (PG), precisely, their glycosaminglycan chains (GAG), most frequently heparan sulphate (HS). Evidences to this, qualitative estimations of the binding/dissociation constants included, are given in many works [18,19,27,29,41,42,43], but see [26] (there also exist data on binding of e.g. penetratin to membranes of liposomes [44]). Here the main role belongs to electrostatic interactions and neutralization of opposite charges in the formed complexes of positively charged PTDs and negatively charged HSs. In early works, in which a direct interaction of PTD with phospholipides of the plasma membrane was supposed, the inverted micelle mechanism was suggested [13], see Fig.2. This model gained reminiscences in recent work [45] where, in a word, a reversing of the membrane (displacement of phosphatidylserine to outer surface of the cell membrane) resulted from TAT PTD transduction was registered. Allusions to possible formation of a membrane pore are relatively rare since the latter of a needed size presumably would be fatal for the cell.

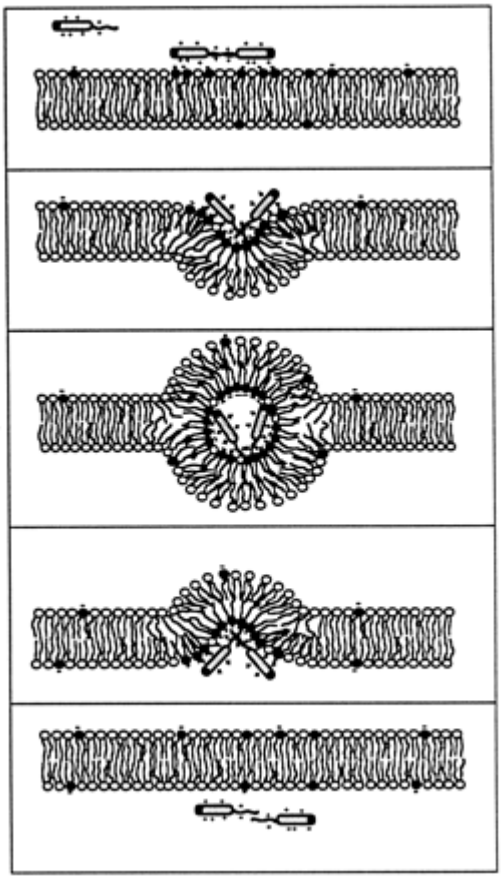


Fig.2. The peptide, represented as a dimer, recruits negatively charged phospholipids (filled circles) and induces the formation of an inverted micelle. The hydrophilic cavity of the micelle accommodates the peptide and, possibly, sequences attached to it that can subsequently be released in the cytoplasmic compartment [13].

Kinetic analysis of the translocation process is hindered by experiment difficulties, that is why data on kinetics are scanty and dispersive. The general conclusion can be reduced to a relatively fast internalization of peptides detected in the plasma membrane and cytosol sometimes even in several minutes, and in 20÷60 min – often in the nuclear membrane/nucleus. To the moment, as is justly mentioned by Zorko and Langel [6], it is reasonable to consider a phenomenological kinetic scheme like shown in Fig.3 with all its stages being sufficiently justified. From their point of view, however, even this scheme is too complicated and needs further simplification. Naturally, such a kinetic analysis could yield indicative results only for revealing the translocation mechanism. But, astonishingly, even this necessary step is practically

absent in the whole body of literature in the field, where attention is paid mostly to qualitative or physiological/medical aspects of experimental results. Although it goes without saying that targeted application of the phenomenon to drug delivery and clinical therapy will be impossible without full comprehension of its physicochemical mechanism.

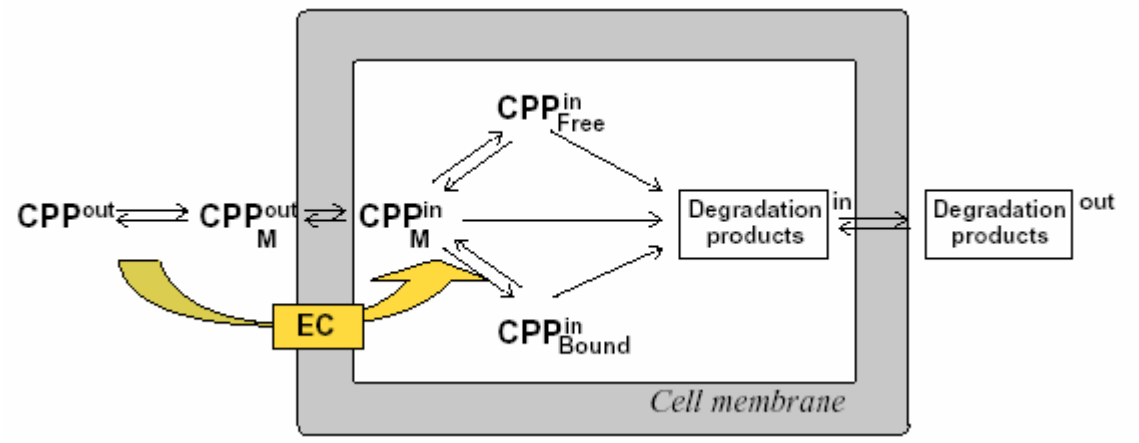


Fig.3. Simplified kinetic scheme for CPP internalization. Indices *in* and *out* represent the portions of CPP or its degradation products inside and outside the cell, respectively; *M* denotes membrane bound CPP; *Free* means internalized but non-bound CPP (e.g. in cytosol); *Bound* means the fraction of CPP that is interacting with inner cell structures (intracellular membranes, proteins, etc); *Degradation* products result from proteolytic cleavage of the CPP in the cell; *EC* denotes endocytosis [6].

3. Planned ways and methods of further investigation

Resuming the previous section, one can note the following. First works done on cell-penetrating peptides (CPPs) can be classified as qualitative observation of the peptide uptake by cells. Leaving aside the questions of adequate experimental protocols it is possible to postulate that the effective uptake of pure CPPs or different cargoes linked to CPPs is now established with no doubt. However, apparent uptake of CPPs does not mean that the peptides cross the membrane in a “mysterious” way - the uptake can be caused by the well known endocytic pathway. There is a significant body of studies that confirm either endocytic internalization or direct membrane crossing by CPPs, but they can be hardly considered as absolutely convincing. It was postulated recently that several most widely studied CPPs such as penetratin, TAT and synthetic oligoarginines are internalized concurrently by both endocytosis and direct membrane crossing. It was established that the binding of CPPs to the membranes is non-specific and is necessary for both direct crossing and endocytosis.

The endocytic uptake of CPPs is of little interest from the point of view of the current project. In this case the peptide resides at the same extracellular compartment and the very complex energy-dependent process of vesicle formation facilitates the uptake. In contrast, direct translocation of CPPs through the membrane is of great interest. During this process the soluble CPP molecule manages to cross the hydrophobic core of the membrane. This requires some unknown physical mechanisms making it possible to overcome a prohibitively high energy barrier of such translocation. That is why we henceforth focus on direct translocation of CPPs through the membrane.

General physical picture of translocation. In general, the translocation mechanism can be divided into several stages regardless of its molecular details. The first stage is the binding of the CPP molecule to the membrane surface. This process is energetically favorable. The next stage is translocation through the hydrophobic core of the membrane. This stage most likely involves the crossing of a substantial energy barrier or several such barriers. After this stage, the peptide appears bound on the internal side of the membrane. The last stage is the release of the peptide to the cytosol that is again energetically unfavorable because of the binding energy. The translocation is spontaneous and does not require external energy sources like ATP, thus the whole process can be viewed as a thermo-activated diffusion process. It is possible to introduce the *translocation coordinate x*, which can be roughly associated with the position of the CPP in the membrane. A hypothetic energy profile along the translocation coordinate consists of two energy minima that correspond to the CPP bound to the external or internal sides of the membrane. These minima are divided by the energy barrier that corresponds to the peptide location in the hydrophobic core of the membrane. The exact shape of the energy profile in this region depends strongly on molecular details of translocation. Several possible variants are discussed below.

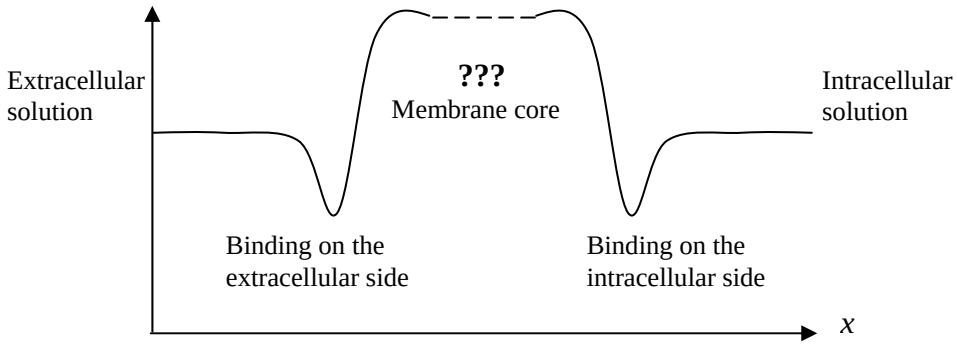


Fig.4.

As noted above, there are several models of translocation mechanisms in the literature (inverted micelle model, carpet model, etc). It is also possible that the single peptide intercalates

between the lipids and cross the membrane without forming specific pore-like structures. Let us consider these cases in more details.

Inverted micelle model and carpet model. These imply that the lipid bilayer is reorganized in such a way that no contact between the peptide and the hydrophobic lipid tails is necessary during translocation. In fact the peptide is coated by the lipid heads.

The inverted micelle model implies that the peptide molecule induces the formation of a local depression on the surface of the membrane. This can be accomplished by strong binding of the peptide to the head groups of several lipids. The depression can then close itself into an inverted micelle. If this micelle is opened to the internal solution, the peptide is transferred through the membrane. However, the feasibility of this mechanism is questionable: the minimal size of inverted micelle is quite large, leading to severe distortion of the outer leaflets of the bilayer; thus the structure can hardly be stable.

The carpet model implies that a number of peptide molecules cover the surface of the membrane as a patch of a carpet. Interactions between the peptides and lipids destabilize the planar structure of the bilayer and induce the formation of a pore, covered by the lipid molecules and peptides. Such a pore does not expose hydrophobic lipid tails to the water environment – it is covered by the lipid head groups with bound peptides.

The barrel-stave model, in contrast, postulates that the peptides themselves form the pore walls. The leaflets of the lipid bilayer in this case are not bent at all. The peptide molecules are arranged like a barrel with the hydrophobic outer surface contacting the lipids and the hydrophilic inner surface that forms the pore [46].

All the mentioned structures should be meta-stable with the lifetime comparable to the translocation time. However, formation of such structures is energetically unfavorable because it requires a significant perturbation of the bilayer. As a result, the energy profile for these models should be like that in Fig.5:

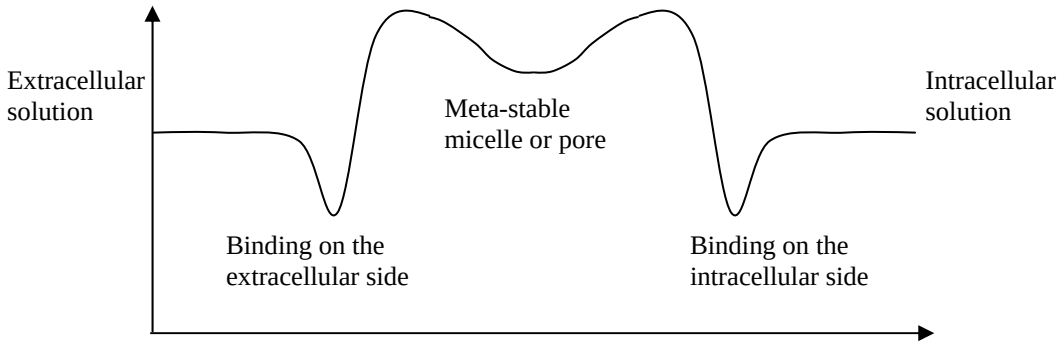


Fig.5.

Direct translocation. If the peptide intercalates between the hydrophobic tails of the lipids and crosses the membrane directly without formation of a specific membrane structure then the energy of peptide inside the membrane core should be the highest and the energy profile will have a single barrier (Fig.6).

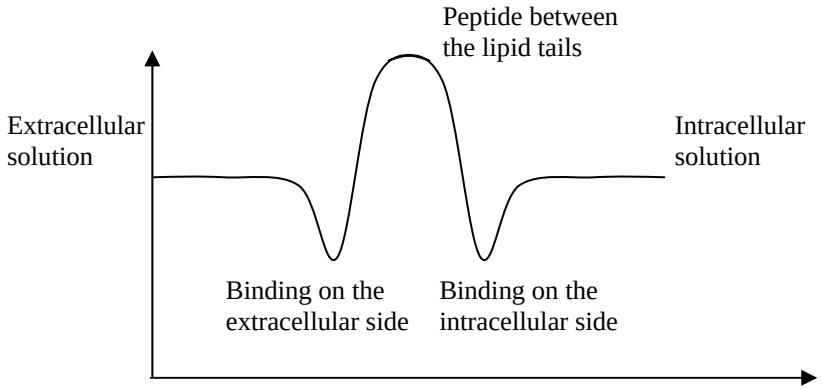


Fig.6.

Quantitative parameters of binding. Binding of the CPP to the membrane surface is the crucial stage of translocation, presenting in any translocation model. Recent advances in experimental techniques allow one to determine the binding properties quantitatively. To our knowledge, most of recent quantitative data relate to penetratin or Tat PTD.

The binding parameters can be evaluated using the binding isoterms. The only parameters which can be directly extracted from the experiment is the membrane-bound peptide to lipid molar ratio and apparent dissociation constant of the peptide. Using the model of Langmuir adsorption it is possible to extract the binding energy E_b by fitting the binding isoterms. A very interesting attempt to extract different components of the binding energy was published recently [44]. The authors use the Gouy-Chapman theory that allows them to estimate the electrostatic and hydrophobic components of the binding energy. In addition it is possible to extract such quantities as the local concentration of the peptide near the membrane, membrane surface charge density, membrane surface potential and effective peptide charge. These can be used in modeling. They are especially useful because the membrane and peptide charges can be adapted to represent not only the outer solution, but also the cell interior, which is necessary to calculate the binding to the cytosolic side of the membrane.

Kinetics of internalization. As noted above, internalization of penetratin and other CPPs includes two mechanisms – direct penetration and endocytosis. Zaro and Shen [15] published the experimental data that allows one to compare the importance of these pathways quantitatively. They measured the time-resolved concentration of the internalized peptide in the vesicular fraction and in the cytosol separately. The authors did not calculate any kinetic constants, but

their data are perfectly suited for such an analysis. The only serious limitation is a very small number of points on the curves, which can make the fitting ambiguous. Let us consider the time-dependent concentrations of the peptide in the vesicles, $V(t)$, and in the cytosol, $C(t)$, measurable experimentally. The peptide concentration in the outer solution is O (Fig. 7).

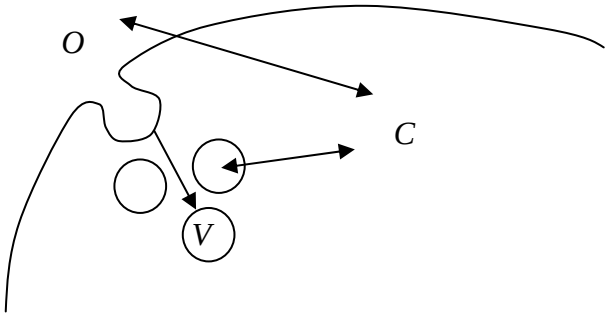


Fig.7.

The processes of direct penetration and vesicle formation can be approximated by single-stage first-order reactions. One can define the kinetic constants of peptide exchange between the three compartments (cytosol, vesicles and outer solution), k_{cv}, k_{vc}, k_{co}, k_{oc}, k_{vo}, k_{ov}.

The following balance equations describe the behavior of the system:

$$\dot{C} = k_{vc}V + k_{oc}O - (k_{cv} + k_{co})C$$
$$\dot{V} = k_{cv}C + k_{ov}O - (k_{vc} + k_{vo})V$$

This set contains six empirical kinetic constants. The quality of experimental data does not allow us to determine all the six constants by direct fitting the curves. That is why further simplifications are needed. The endocytosis is an energy-dependent process that can be considered as irreversible, thus the vesicles can hardly bring their cargo back to the outer solutions and we can set $k_{vo} = 0$. Next, it is possible to assume that the direct permeation through the membrane of the vesicle is essentially the same as through the outer membrane (in reality this can be modified by the acidic conditions inside the vesicles). In this case $k_{vc} = k_{oc} = k_{in}$; $k_{cv} = k_{co} = k_{out}$. The remaining three constants can be determined by fitting the experimental data.

Determining the energy profile of direct translocation. Analysis of the kinetics allows one to obtain reliable quantitative kinetic constants k_{in} and k_{out} that describe the inward and outward penetration through the membrane. According to the Kramers theory of thermo-activated reactions, one can write:

$$k_{in} = A\exp\left(-E_{in}/k_BT\right)$$
$$k_{out} = A\exp\left(-E_{out}/k_BT\right)$$

where E_{in} and E_{out} are the heights of the effective energy barriers of penetration, A is the Kramers factor. The effective energy barrier depends on the heights of actual barriers along the translocation coordinate. If there are no significantly populated intermediate states on the translocation pathway, then the effective barrier corresponds to the rate-determining highest barrier along the translocation coordinate.

The heights of the other energy barriers cannot be extracted using kinetic data only. In order to do this the whole body of experimental data should be involved. The following scheme (Fig.8) shows which kind of data can be used to estimate the heights of various barriers:

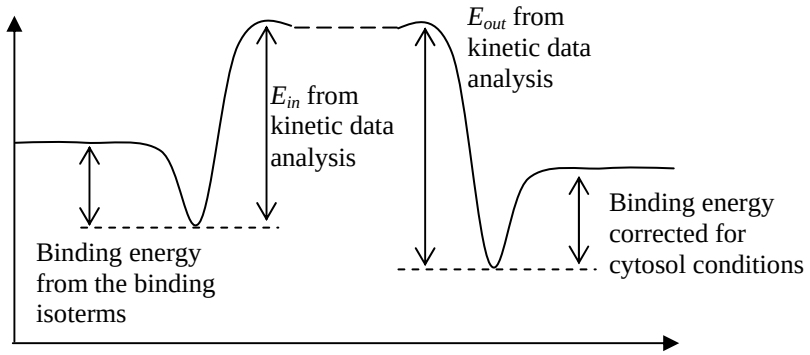


Fig.8.

Theoretical estimate of the rate-determining energy barrier. The values of E_{in} and E_{out} that are most probably the rate-determining barriers heights obtained from experimental data become invalid if different peptides or different membrane composition are considered. Thus, independent theoretical estimates are necessary. One of the possible approaches is direct molecular dynamics calculations of the molecular membrane models corresponding to various penetration mechanisms. This approach can discriminate between possible mechanisms and determine the most probable mechanism for a particular membrane composition, penetrating peptide and cargo. However, such calculations require the modeling of quite large membrane patches for quite long periods of time that presumes the usage of supercomputers containing not less than 100 processors for several months.

Alternatively, simplified theoretical models can be developed for each penetration mechanism. This approach does not require extensive calculations but is error-prone because of a number of empirical data incorporated into such models. The lack of quantitative data does not allow one to extract the empirical constants with proper accuracy for a wide range of membranes

and penetrating peptides. Further progress needs experiments closely coupled with theoretical studies and directly targeted on the verification of specific theoretical models.

General approach and formalism. To construct a quantitative model of the phenomenon we will consider the translocation process as a diffusion process with allowance for structural properties of the (CPP + cargo) complex, its interaction with the membrane, and the structural and dynamical properties of the membrane. Here we outline the possible way of performing this task.

The diffusion process of crossing the membrane by the CPP is conditioned by fluctuations in positions of structural components of the CPP and membrane. This implies the two most applicable methods of description: Langevin stochastic equations or corresponding Fokker-Planck equations for the multidimensional distribution function of system's variables. Let $V_{\mathrm{CPP}}(\boldsymbol{x})$ be potential expressed in terms of the CPP's structural variables $\boldsymbol{x} = \left(x_1, x_2 ... x_N\right)$ which describe the relative motion of CPP's structural components and motion of the CPP as a whole. Next, let $V_M(\boldsymbol{x})$ be potential determining the dynamics of structural components $\boldsymbol{y} = \left(y_1, y_2 ... y_M\right)$ of the membrane, and $W(\boldsymbol{x}, \boldsymbol{y})$ stands for the interaction between the CPP and membrane. Then the Langevin stochastic equations describing both dynamical and diffusive processes in the system read:

$$q_i \frac{dx_i}{dt} = -\frac{\partial U(\boldsymbol{x}, \boldsymbol{y})}{\partial x_i} + \sqrt{2k_B T q_i}\,\xi_i(t), \qquad \left\langle \xi_i(t)\xi_j(t') \right\rangle = \delta_{ij}\delta(t-t')$$

$$q_j \frac{dy_j}{dt} = -\frac{\partial U(\boldsymbol{x}, \boldsymbol{y})}{\partial y_j} + \sqrt{2k_B T q_j}\,\xi_j(t), \qquad i = \{1, ...N\}, \qquad j = \{1, ...M\},$$

where $U(\boldsymbol{x}, \boldsymbol{y}) = V_{CPP}(\boldsymbol{x}) + V_M(\boldsymbol{y}) + W(\boldsymbol{x}, \boldsymbol{y})$ is the total potential energy of the system (CPP + M) and q_i is the friction coefficient for *i*-th coordinate.

The corresponding Fokker-Planck equation for the joint distribution function $P(\boldsymbol{x}, \boldsymbol{y}, t)$ reads [47-48]:

$$\left\{ \frac{\partial}{\partial t} - \sum_{i=1}^{N} D_{x_i} \frac{\partial}{\partial x_i} \left[\frac{1}{k_B T} \frac{\partial U(\boldsymbol{x}, \boldsymbol{y})}{\partial x_i} + \frac{\partial}{\partial x_i} \right] - \sum_{j=1}^{M} D_{y_j} \frac{\partial}{\partial y_j} \left[\frac{1}{k_B T} \frac{\partial U(\boldsymbol{x}, \boldsymbol{y})}{\partial y_j} + \frac{\partial}{\partial y_j} \right] \right\} P(\boldsymbol{x}, \boldsymbol{y}, t) = 0$$

where the diffusion coefficients $D_{x_i} = \frac{k_B T}{q_i}$ and $D_{y_j} = \frac{k_B T}{q_j}$ relate to structural variables of the CPP and membrane, respectively.

In both cases, in view of the real number of variables, to directly follow the equations above, in either numerical or analytical calculations, is a daunting task. Here extremely helpful in necessary simplifications are the methods of adiabatic elimination of fast variables which we successfully applied earlier for description of biomolecular systems [49-50]. We also intend using the Brownian dynamics methods for the most important slow system's variables (to be developed at the next stage of the project).

Construction of a dynamical model of the CPP is one of the main auxiliary tasks. Such a model, on the one hand, should be simple enough and accessible computationally. On the other hand, it should reflect specific features of a given CPP that determine its penetration capability. Preliminary analysis shows that such a model can be based on the notion of the CPP as a system of several blocks mutually interacting by potential $V_{\mathrm{CPP}}(\boldsymbol{x})$. The relative movements of the blocks will be identified by degrees of freedom $\boldsymbol{x}$, and the corresponding parameters of the latter will determine the object specificity. It is supposed that a block will be relatively large (at a level of a residue or larger) to decrease the number of degrees of freedom to reasonable extent but without loss of the CPP specificity. We consider the use of the coarse-grained knowledge based potentials (see e.g. [511]) for constructing $V_{\mathrm{CPP}}(\boldsymbol{x})$ as acceptable in modeling.

Not less important is the way of modeling the membrane. Here we also intend holding the method of reasonable sufficiency. If the coarse-grained model is exploited in modeling the CPP, then the continuum approximation will be sufficient for modeling the membrane since tracing the behaviour of individual lipid molecules is not necessary. The following model of the membrane can be suggested: the hydrophobic central part of the membrane (tails of lipid molecules) is represented by an infinite homogeneous and isotropic dielectric layer of finite thickness d_{core} with a small dielectric constant ($\varepsilon = 2 \div 3$) taken from experiment. This layer is covered on both sides with two dielectric layers of thickness d_{polar} that correspond to the domains of polar heads of lipid molecules. These latter layers have its dielectric constant as typical for partially ordered polar medium ($50 < \varepsilon < 80$). The polar layers possess bulk and/or surface charges representing the charges of real lipid heads, The charge distribution can be rather inhomogeneous in the direction normal to the bilayer but isotropic in its plane. Such a model allows us to calculate the potential profiles for each type of penetrating particles (structural blocks inside the peptides). It is possible to vary the lipid composition of the membrane and asymmetry of the internal and external monolayers by means of varying the dielectric properties and charge distributions of the polar layers. At the same time the model admits sufficiently simple ways of taking into account the interaction of different CPPs with the membrane.

The simplest can be carried out within the constant potential approximation, $\vartheta(\boldsymbol{x}) = W(\boldsymbol{x}, \boldsymbol{y_0})$. Then the positions of the membrane structural groups are supposed to be fixed

at given mean values $\boldsymbol{y_0}$, and $\vartheta(\boldsymbol{x})$ determines the interaction of each of movable elements of the CPP with the membrane. In this case the problem is reduced to the study of the over-crawling of a non-rigid object of the given structure containing components of different mobility through a rigid pore (cf. e.g. [52-54]). More sophisticated description allows for self-consistent dynamics of the (CPP + membrane) system. It is the version of the theory that could turn out to be most suitable for describing the striking features of protein transduction.

Литература

1. M. Green, P. Loewenstein. Autonomous functional domains of chemically synthesized human immunodeficiency virus tat transactivator protein. *Cell* **55**, 1179-1188 (1988).
2. A. Frankel, C. Pabo. Cellular uptake of the tat protein from human immunodeficiency virus. *Cell* **55**, 1189-1193 (1988).
3. D. Mann, A. Frankel. Endocytosis and targeting of exogeneous HIV-1 Tat protein. *EMBO J.* **10**, 1733-1739 (1991).
4. S. Fawell, J. Seery, Y. Daikh, C. Moore, L. Chen, B. Pepinsky, J. Barsoum. Tat-mediated delivery of heterologous proteins into cells. *PNAS USA* **91**, 664-668 (1994).
5. S. Schwarze, A. Ho, A. Vocero-Akbani, S. Dowdy. In vivo protein transduction: Delivery of a biologically active protein into the mouse. *Science* **285**, 1569-1572 (1999).
6. M. Zorko, U. Langel. Cell-penetrating peptides: mechanisms and kinetics of cargo delivery. *Adv. Drug Deliv. Rev.* **57**, 529-545 (2005).
7. New tools enable gene delivery. *The Scientist* **14** [24], 30 (2000).
8. Getting proteins into cells. *The Scientist* **16** [7], 38 (2002).
9. S. Moghimi, A. Hunter, J. Murray. Nanomedicine: current status and future prospects. *FASEB J.* **19**, 311-330 (2005).
10. M. Lindgren, M. Hallbrink, A. Prochiantz, U. Langel. Cell-penetrating peptides. *Trends Pharmacol. Sci.* **21**, 99-103 (2000).
11. E. Eriksdottir, H. Myrberg, M. Hansen, U. Langel. Cellular uptake of cell-penetrating peptides. *Drug Design Rev. Online* **1**, 161-173 (2004).
12. P. Wender, D. Mitchell, K. Pattabiraman, E. Pelkey, L.Steinman, J.Rothbard. The design, synthesis and evaluation of molecules that enable or enhance cellular uptake: Peptoid molecular transporters. *PNAS USA* **97,** 13003-13008 (2000).
13. D. Derossi, S. Calvet, A. Trembleau, A. Brunissen, G. Chassaing, A. Prochiantz. Cell internalization of the third helix of the Antennapedia homeodomain is receptor-independent. *J. Biol. Chem.* **271**, 18188-18193 (1996).
14. G. Drin, M. Mazel, P. Clair, D. Mathieu, M. Kaczorek, J, Temsamani. Physico-chemical requirements for cellular uptake of pAntp peptide. Role of lipid-binding affinity. *Eur. J. Biochem.* **268**, 1304-1314 (2001).
15. J. Zaro, W.-C. Shen. Qualitative comparison of membrane transduction and endocytosis of oligopeptides. *Biochim. Biophys. Res. Comm.* **307**, 241-247 (2003).
16. S. Futaki, T. Suzuki, W. Ohashi, T. Yagami, S. Tanaka, K. Ueda, Y, Sugiura. Arginine-rich peptides. *J. Biol. Chem.* **276**, 5836-5840 (2001).

17. T. Suzuki, S. Futaki, M. Niwa, S. Tanaka, K. Ueda, Y, Sugiura. Possible existence of common internalization mechanisms among arginine-rich peptides. *J. Biol. Chem.* **277**, 2437-2443 (2002).
18. C. Mai, H, Shen, S. Watkins, T.Cheng, P. Robbins. Efficiency of protein transduction is cell type-dependent and is enhanced by dextran-sulfate. *J. Biol. Chem.* **277**, 30208-30218 (2002).
19. M. Belting. Heparan sulfate proteoglycan as a plasma membrane carrier. *Trends Biochim. Sci.* **28**, 145-151 (2003).
20. A. Elmquist, M. Lindgren, T. Bartfai, U. Langel. VE-cadherin-derived cell-penetratin peptide pVEC with carrier functions. *Exp. Cell Res.* **269**, 237-244 (2001).
21. M. Pooga, C. Kut, M. Kihlmark, M. Hallbrink, S. Fernaeus, R. Raid, T. Land, E. Hallberg, T. Bartfai, U. Langel. Cellular translocation of proteins by transportan. *FASEB J.* **15**, U304-316 (2001).
22. V. Torchilin, R. Rammohan, V. Weissig, T. Levchenko. TAT peptide on the surface of liposomes affords their efficient intracellular delivery even at low temperature and in the presence of metabolic inhibitors. *PNAS USA* **98,** 8786-8791 (2001).
23. V. Torchilin, T. Levchenko, R. Rammohan, N. Volodina, B. Papahadjopoulos-Sternberg, G. D'Souza. Cell transfection in vitro and in vivo with nontoxic TAT peptide-liposome-DNA complexes. *PNAS USA* **100,** 1972-1977 (2003).
24. D. Derossi, A. Joliot, G. Chassaing, A. Prochiantz. The third helix of the Antennapedia homeodomain translocates through biological membranes. *J. Biol. Chem.* **269**, 10444-10450 (1994).
25. E. Vives, P. Brodin, B. Lebleu. A truncated HIV-1 Tat protein basic domain rapidly translocates through the plasma membrane and accumulates in the cell nucleus. *J. Biol. Chem.* **272**, 16010-16017 (1997).
26. M. Silhol, M. Tyagi, M. Giacca, B. Lebleu, E.Vives. Different mechanisms for cellular internalization of the HIV-1 Tat-derived cell penetrating peptide and recombinant proteins fused to Tat. *Eur. J. Biochem.* **269** 494-501 (2002).
27. S. Sandgren, F. Cheng, M. Belting. Nuclear target of macromolecular polyanions by an HIV-Tat derived peptide. *J. Biol. Chem.* **277**, 38877-38883 (2002).
28. J. Richard, K. Melikov, E. Vives, C. Ramos, B. Verbeure, M. Gait, L. Chernomordik, B. Lebleu. Cell-penetrating peptides. A reevaluation of the mechanism of cellular uptake. *J. Biol. Chem.* **278**, 585-590 (2003).
29. A. Fittipaldi, A. Ferrari, M. Zoppe, C. Arcangeli, V. Pellegrini, F. Beltram, M. Giacca. Cell membrane lipid rafts mediate caveolar endocytosis of HIV-1 Tat fusion proteins. *J. Biol. Chem.* **278,** 34141-34149 (2003).

30. T. Potocky, A. Menon, S. Gellman. Cytoplasmic and nuclear delivery of a Tat-derived peptide and a β-peptide after endocytic uptake into HeLa cells. *J. Biol. Chem.* **278,** 50188-50194 (2003).
31. S. Console, C. Marty, C. Garcia-Echeverria, R. Schwendener, K. Ballmer-Hofer. Antennapedia and HIV transactivator of transcription (TAT) "protein transduction domains" promote endocytosis of high molecular weight cargo upon binding to cell surface glycosaminglycans. *J. Biol. Chem.* **278,** 35109-35114 (2003).
32. I. Ignatovich, E. Dirzhe, A. Pavlotskaya, B. Akinfiev, S. Burov, S. Orlov, A. Perevozchikov. Complexes of plasmid DNA with basic domain 47-57 of the HIV-1 Tat protein are transferred to mammalian cells by endocytosis-mediated pathways. *J. Biol. Chem.* **278,** 42625-42636 (2003).
33. P. Thoren, P. Persson, P. Isakson, M. Goksor, A. Onfelt, B. Norden. Uptake of analogs of penetratin, Tat_{48-60} and oligoarginine in live cells. *Biochim. Biophys. Res. Comm.* **307,** 100-107 (2003).
34. R. Veach, D. Liu, S. Yao, Y, Chen, X. Liu, S. Downs. Receptor/transporter-independent targeting of functional peptides across the plasma membrane. *J. Biol. Phys.* **279,** 11425-11431(2004).
35. B. Christiaens, J. Grooten, M. Reusens, A. Joliot, M. Goethals, J. Vandekerckhove, A. Prochiantz, M. Rosseneu. Membrane interaction and cellular internalization of penetratin peptides. *Eur. J. Biochem.* **271** 1187-1197 (2004).
36. G. Drin, S. Cottin, E. Blanc, A. Rees, J. Yemsamani. Studies of the internalization mechanism of cationic cell-penetrating peptides. *J. Biol. Chem.* **278**, 31192-31201 (2003).
37. R. Fisher, K. Kohler, M. Fotin-Mleczek, R. Brock. A stepwise dissection of the intracellular fate of cationic cell-penetrating peptides. *J. Biol. Chem.* **279**, 12625-12635 (2004).
38. S. Sandgren, A. Wittrup, F. Cheng, M. Jonsson, E. Eklund, S. Busch, M. Belting. The human antimicrobal peptide LL-37 transfers extracellular DNA plasmid to the nuclear compartment of mammalian cells via lipid rafts and proteoglycan-dependent endocytosis. *J. Biol. Chem.* **279**, 17951-17956 (2004).
39. L. Hyndman, J, Lemoine, L. Huang, D. Porteous, A. Boyd, X. Man. HIV-1 Tat protein transduction domain peptide facilitates gene transfer in combination with cationic liposomes. *J. Control. Release* **99**, 435-444 (2004).
40. J. Wadia, R. Stan, S. Dowdy. Transducible TAT-HA fusogenic peptide enhances escape of TAT-fusion proteins after lipid raft macropinocytosis. *Nat. Med.* **10**, 310-315 (2004).

41. M. Tyagi, M. Rusnati, M. Presta, M. Giacca. Internalization of HIV-1 Tat requires cell surface heparan sulfate proteoglycans. *J. Biol. Chem.* **276**, 3254-3261 (2001).
42. S. Hakansson, M. Caffrey. Structural and dynamical properties of the HIV-1 Tat transduction domain in the free and heparan-bound states. *Biochemistry* **42**, 8999-9006 (2003).
43. A. Ziegler, J. Seelig. Internalization of the protein transduction domain of HIV-1 TAT with heparan sulphate: Binding mechanism and thermodynamical parameters. *Biophys. J.* **86**, 254-263 (2004).
44. D. Persson, P. Thoren, M. Herner, P. Lincoln, B. Norden. Application of a novel analysis to measure the binding of the membrane-translocating peptide penetratin to negatively charged liposomes. *Biochemistry* **42**, 421-429 (2003).
45. V. Moore, R. Payne. Transactivator of transduction fusion protein transduction causes membrane inversion. *J. Biol. Chem.* **279**, 32541-32544 (2004).
46. P. Lundberg, U. Langel. A brief introduction to cell-penetrating peptides. *J. Mol. Recogn.* **16,** 227-233 (2003).
47. N. van Kampen. Stochasic Processes in Physics and Chemistry. *North-Holland*, Amsterdam, New York, Oxford (1981).
48. C. Gardiner. Handbook of Stochastic Methods. *Springer-Verlag*, Berlin, Heidelberg, New York, Tokyo (1985).
49. L. Christophorov, A. Holzwarth, V. Kharkyanen, F. van Mourik. Structure-function self-organization in nonequilibrium macromolecular systems. *Chem. Phys*. **256**, 45-60 (2000).
50. A. Goushcha, V. Kharkyanen, G. Scott and A. Holzwarth. Self-regulation phenomena in bacterial reaction centers. 1. General Theory. *Biophys. J*. **79**, 1237-1252 (2000).
51. C. Zhang, S. Liu, H. Zhou, Y. Zhou. An accurate residue-level potential of mean force for folding and binding based on the distance-scaled ideal gas reference state. *Protein Sci.* **13**, 400-411 (2004).
52. E. Slonkina, A. Kolomeisky. Polymer translocation through a long nanopore. *J. Chem. Phys.* **118**, 7112-7118 (2003).
53. A. Kolomeisky, A. Drzewinski. Polymer dynamics in repton model at large fields. *J. Chem. Phys.* **120**, 7784-7791 (2004).
54. I.-C. Yeh, G. Hummer. Nucleic acid transport through carbon nanotube membranes. *PNAS USA* **101**, 12177-12182 (2004).

Summary

During the last decade the protein transduction domains or, more generally, cell-penetrating peptides (CPPs) are being of special interest as a promising non-invasive means of delivery of macromolecular components and whole macromolecules into cells and cellular nuclei. Although extensive experimental data have been reported in the literature, they are highly diverse and often contradictory, starting from the very possibility of CPPs' transduction and especially the translocation mechanisms. We review the current state of the art in the field and outline possible physical approaches to the modeling and quantitative analysis of the phenomenon of transduction.

EOARD Partner Project 211 – Deliverable 2

1. Introduction

As was noted in Deliverable 1, from the available amount of data it is hard to distinguish more or less special common features in the primary/secondary structure of the cell-penetrating peptides (CPPs) except the fact that all of them are positively charged. This suggests that the modeling of the CPP at the specific structure level can be preceded by a consideration in which the CPP is represented by a simple dielectrical stereoobject (sphere/ellipsoid, cylinder, etc) with some distribution of its positive charge. The electrostatic interaction with the negatively charged surface of the lipid bilayer membrane will naturally course approaching of the CPP and membrane, accompanied with considerable changes in the structure and geometry of the latter, up to formation of some complex whose character will allow us to evaluate the transduction possibility.

Obviously, in such problem formulation the leading part is played by the method of modeling of the membrane structure. This can be done within the mechanical continuous models by the methods of thin film elasticity theory, etc (see e.g. [1]). The main parameters entering these models are material constants like coefficients of surface tension, bending elasticity, spontaneous curvature, so on. Although such approaches made a good showing in some problems of the rearrangements of lipid aggregates (e.g. bilayer-micelle transitions), it should be noted that rigorously they are applicable for describing small membrane deformations only. In the presence of domains of essentially different lipid packing they can produce, at their best, only qualitatively correct results.

Another possibility is to use a semi-microscopic approach based on the so-called "molecular lipid models" originating from works of Israelachvili *et al* on lipid self-assembly (see e.g. [2-4] and "the opposing forces model" (OFM) proposed therein); in recent years it is intensively developed, in particular, by Ben-Shaul and co-authors [5-9]. In this approach the lipids are treated as some stereoobjects (cones, cylinders, etc) able in certain limits to change their shape, depending on the type of packing in a lipid aggregate, without changing their volume (fluid model of constant density). Within such models all the main forces determining the aggregate shape can be represented by a small number of relatively simple contributions into the free energy per lipid molecule.

Thus, the modeling strategy at the present stage looks as follows. The free energy of the system (peptide + membrane) is minimized by variational methods. With this, both the free energy profile and system geometry (changing as the peptide approaches to the lipid bilayer, up to critical geometrical changes indicating the possibility of transduction) can be found. Typical

realistic values of parameters of the bilayer representing an artificial lipid membrane, as well as those of the peptide, are more or less known. The influence of the electrolyte surrounding the membrane and peptide can be considered in a simple Debye model or in a more detailed way involving the Poisson-Boltzmann equation [7, 10-12].

Below we briefly recall the necessary knowledge of the lipid membranes and the opposing forces model and then expound our modernization of this model and the method of calculation of the free energy which will be exploited at the next stage of the project.

2. Lipid aggregates and the extended opposing forces model

When trying to create molecular models of lipid aggregates self-assembly it has turned out that rather good results (although initially at a qualitative level of prediction of possible structures only) could be obtained within a relatively rough phenomenological approach using mainly the lipid geometrical properties only. For example, one can compose the dimensionless packing parameter $p = \upsilon / a_0 l_c$ whose value immediately indicates the type of a possible lipid aggregate.

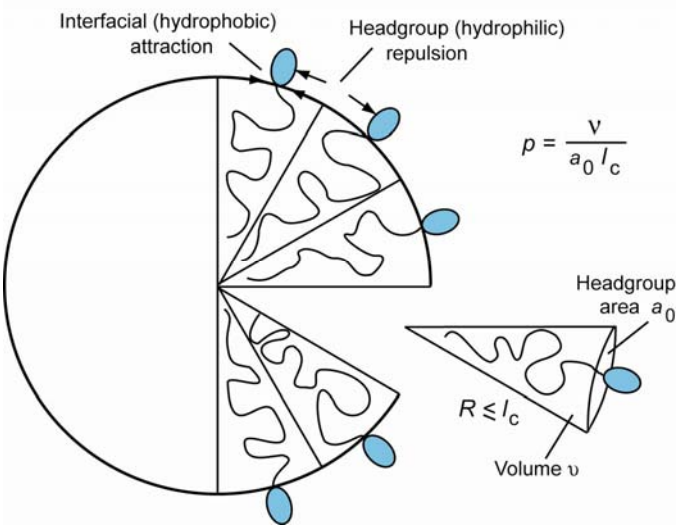
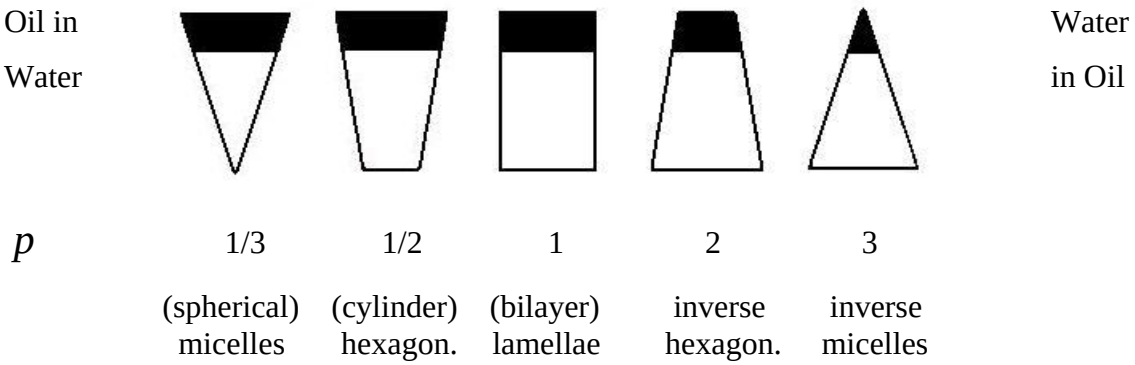


Fig.1. The surface and bulk intermolecular interactions that define the dimensionless packing parameter, or average molecular "shape factor" p of amphiphiles [4].

Here υ is the volume per lipid molecule, l_c is the maximum possible extension of the flexible hydrocarbon chains, and a_0 is the "optimum" (i.e. corresponding to minimum free energy) headgroup area at the hydrocarbon-water interface, see Fig.1. There exists the so-called "generic sequence" of the main amphiphilic structures and corresponding types of the elementary lipid volumes appearing with p growing:



with numerous more complex intermediate structures [4].

Stability of this or that structure is determined by a subtle balance of the forces on the bilayer surface and in the hydrophobic core. Nevertheless, to a rather good extent, this balance can be represented by the following additive contributions into the free energy per lipid, f. The simplest is the classical expression [2,3]

$$f = f_s + f_h, \tag{1}$$

describing the interplay of the two opposing tendencies: to minimize the hydrocarbon-water interface surface due to hydrophobic interactions (f_s) and to drive the hydrophilic headgroups to move away from each other, maximizing their contact with water (f_h). In the opposing forces model the contribution f_s is expressed through the surface tension coefficient γ:

$$f_s = \gamma a, \tag{2}$$

where a is the area per lipid headgroup, and f_h reads

$$f_h = \frac{B}{a}, \tag{3}$$

being some kind of treating the headgroup interaction within a two-dimensional van der Waals equation of state.[1] The competition of these two forces results in the existence of the so-called

[1] The repulsion expressed by Eq. (3) can reflect not only the elecrostatic interaction of charged headgroups but also other effects like those of excluded volume, etc.

"optimum" headgroup area a_0 that can be easily found by minimization of f (1): $a_0 = \sqrt{B/\gamma}$, so that for small deviations of a from a_0 Eq.(1) can be reduced to

$$f(a) = 2\gamma a_0 + \frac{\gamma}{a_0}\left(a - a_0\right)^2. \qquad (4)$$

The simplest OFM (1-4) is applicable, rigorously speaking, to lipid self-assembly into uniformly packed aggregates. In more complex cases it should be taken into account that, first, the mentioned opposing forces act at somewhat different surfaces (see Fig.1). Accordingly, there is a difference between the headgroup area at the hydrocarbon-water interface a_i and that related to the maximum headgroups interaction surface a_h, with the distance l_h between these surfaces, so that

$$f = \gamma a_i + \frac{B}{a_h}. \qquad (5)$$

The distinction between a_i, a_h and a_0 contributes, in particular, to the lipid layer curvature effects. Besides, there exists the conformational contribution of the lipid tails into the free energy, f_c. Given non-compressibility (constant volume per lipid υ) and uniform density of "lipid liquid" in the hydrophobic core, this contribution can be often written as $f_c = \tau\left(b - l_c\right)^2$ where b is the average length of the lipid chain (detailed calculations of the chain packing in the statistical mean-field theory confirm the validity of such approximation [9,13]). Therefore, the extended OFM is represented by the expression

$$f = f_s + f_h + f_c \equiv \gamma a_i + \frac{B}{a_h} + \tau\left(b - l_c\right)^2. \qquad (6)$$

As shown in e.g. works [6,8], it is quite applicable to the description of rather complex non-uniform lipid aggregates and lipid-protein interactions.

Typical values of the constants entering Eq.(6) are rather well-defined in the literature (see e.g. [4,8,14]). Thus, for a saturated hydrocarbon chain $l_c \le l_{\max} = (0.154 + 0.126n)$ nm, and $\upsilon = (27.4 + 26.9n)\cdot 10^{-3}$ nm^3, where $l_{\max}$ is the fully extended molecular length of the lipid chain containing n alkyl groups. Some of these values used in calculations of amphiphilic hydrocarbon systems in water are listed in Table 1.

Table 1

Surface tension of the hydrocarbon-water interface γ	$0.12 k_B T / \text{Å}^2$, or $50\,\text{mJ/m}^2$ at $T = 300\,\text{K}$
Optimal headgroup area a_0	$50 \cdot 10^{-19}\,\text{m}^2$
Lipid volume υ (a lipid with two C_{18} chains)	$10^{-27}\,\text{m}^3$
Critical chain length l_c $(n = 12 \div 18)$	$10 \div 20\,\text{Å}$
Packing parameter $p = \upsilon / a_0 l_c$ $(n = 18)$	0.8
Lipid tail compressibility τ $(n = 12 \div 16)$	$0.11 \div 0.08\, k_B T / \text{Å}^2$

3. Modification of the extended OFM

For the process under study that suggests considerable changes in membrane geometry we modify the approach described in work [8] and applied there to the description of the structure of the edge of a planar lipid bilayer and calculation of the corresponding line tension. Precisely, in the approximation used below we assume that each lipid monolayer of the membrane is represented by a layer of directed molecules of two-dimensional liquid (smectic liquid crystal). The length and tilt of the molecule director determine the monolayer thickness at a given point. So far, we consider only one type of lipids composing the monolayers and neglect the interaction between the latter. The ends of the lipid tails of both monolayers are situated on the membrane "midline".

Fig.2 shows a schematic model of such a bilayer membrane. According to supposed cylindrical symmetry of the problem, we imply the cylindrical reference frame (R, φ, z) with the peptide situated on the symmetry axis z. In this frame the position of a point is given by radius-vector $\vec{\rho} = \vec{i} R\cos\varphi + \vec{j} R\sin\varphi + \vec{k} z$ where $\vec{i}, \vec{j}, \vec{k}$ are the orts of the cartesian frame (x, y, z):

$$x = R\cos\varphi$$
$$y = R\sin\varphi$$
$$z = z,$$

where $R = x^2 + y^2$ is the distance between the point and z-axis, and φ is the angle between the projection of the radius-vector $\vec{\rho}$ onto plane (x, y) and x-axis.

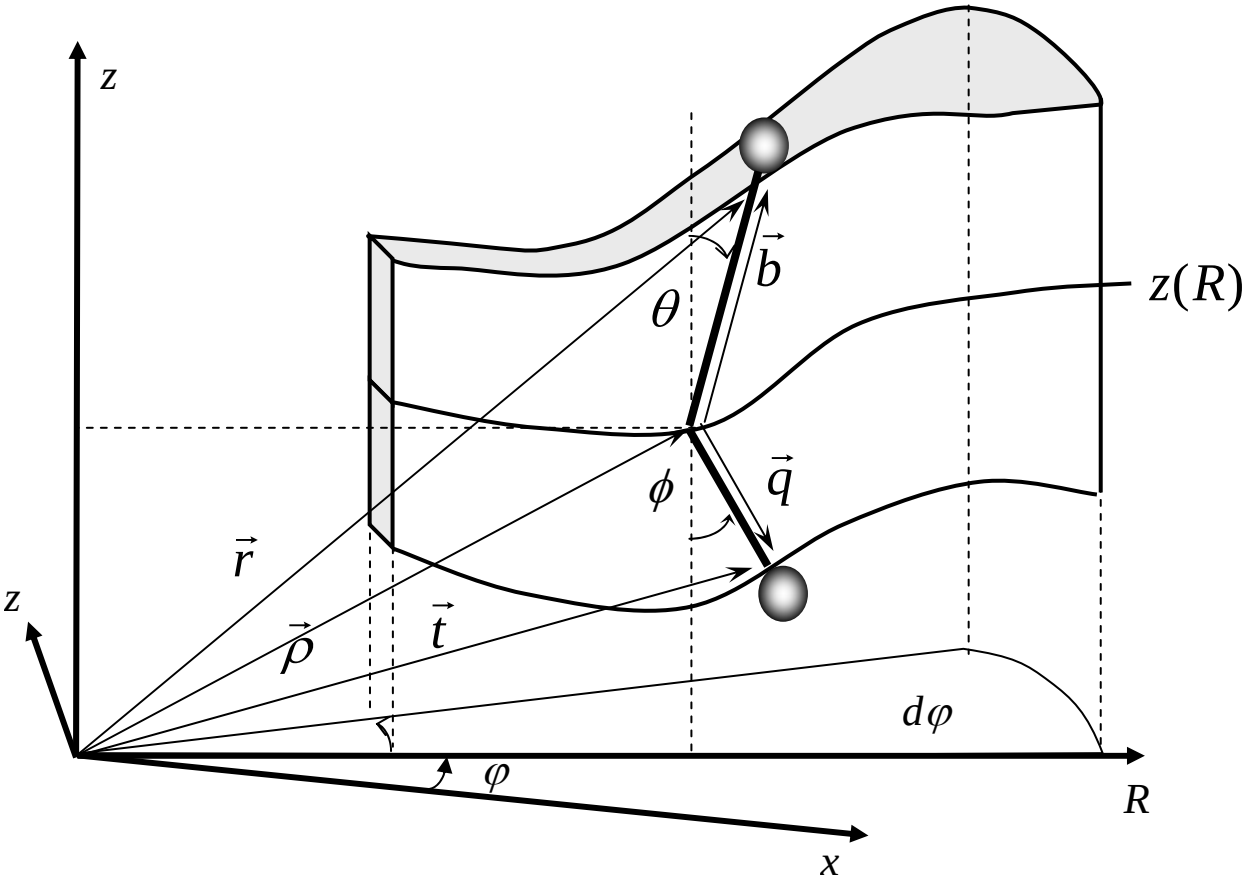


Fig.2.

We now need a proper method of description of membrane geometry changing as the peptide approaches the membrane. As distinct from that used in work [8], we take into account (i) cylindrical symmetry of the system; (ii) nonequivalence of the monolayers and complex shape of the interlayer surface; (iii) the presence of the peptide of a given geometry and charge distribution.

3.1. Geometrical description. As a vector that defines the average position of the inner end of a lipid molecule we take radius-vector $\vec{\rho}$ of a point on the membrane midline (in fact, the set of such points defines the mentioned midline). In the outer ("upper") monolayer, the average position of the molecule with its inner end at point $\vec{\rho}$ is denoted as $\vec{b}(\vec{\rho})$ and the position of the outer end of the hydrocarbon tail (near the headgroup) – as $\vec{r}(\vec{\rho}) = \vec{\rho} + \vec{b}(\vec{\rho})$. Finally, the position of the headgroup center is $\vec{b}(\vec{\rho})\left(1 + l_h / b\right)$, where l_h is the distance between the end of vector $\vec{b}$ and the center of its headgroup. In the inner ("lower") monolayer the corresponding vectors are $\vec{q}(\vec{\rho})$, $\vec{t}(\vec{q}) = \vec{\rho} + \vec{q}(\vec{\rho})$, and $\vec{q}(\vec{\rho})\left(1 + l_h / q\right)$, respectively (see Fig.2). Of course, it is supposed that, in view of the system symmetry, any physical value is φ-independent,

$A(\varphi, R, z) = A(R, z)$, so that the problem can be reduced to a two-dimensional one in plane (R, z).

Suppose that the interlayer midline is described by function $z = z(R)$. Let us parametrize all the necessary functions with respect to R as a parameter. The two-dimensional radius-vector $\vec{\rho}$ is $\vec{\rho} \equiv \vec{\rho}(R) = \vec{e}R + \vec{k}z(R)$ (where $\vec{e} = \vec{R}/R = \vec{i}\cos\varphi + \vec{j}\sin\varphi$), and $d\vec{\rho}(R) = \left(\vec{e} + z'(R)\vec{k}\right)dR$. Here the prime denotes the derivative with respect to R. Consequently, $\vec{b} \equiv \vec{b}(R) = \vec{e}b_R(R) + \vec{k}b_z(R)$ where $b_R = b\sin\theta$, $b_z = b\cos\theta$ are the component of vector $\vec{b}(R)$ and θ is the angle between vector $\vec{b}(R)$ and z-axis (tilt). Then

$$\vec{r}(R) = \vec{\rho}(R) + \vec{b}(R) = (R + b_R)\vec{e} + (z + b_z)\vec{k},$$
$$d\vec{r}(R) = \left[(1 + b_R')\vec{e} + (z' + b_z')\vec{k}\right]dR$$

For the lipids whose inner tails fall into the interval $(R, R + dR)$, the infinitesimal length of the line of the hydrocarbon interface in plane (R, z) reads:

$$\left|d\vec{r}\right| = dR\sqrt{(1 + b_R')^2 + (z' + b_z')^2}$$

in the upper layer and

$$\left|d\vec{t}\right| = dR\sqrt{(1 + q_R')^2 + (z' + q_z')^2}$$

in the lower layer. Then the infinitesimal area on the upper hydrocarbon interface is

$$da_i^{(u)} = (R + b_R)d\varphi\left|d\vec{r}\right| = s_u(R)d\varphi dR, \qquad (7)$$

where

$$s_u(R) = (R + b_R)\sqrt{(1 + b_R')^2 + (z' + b_z')^2}, \qquad (8)$$

and on the lower interface

$$da_i^{(l)} = (R + q_R)d\varphi \left|d\vec{t}\right| = s_l(R)d\varphi dR,$$

where

$$s_l(R) = \left(R + q_R\right)\sqrt{(1+q_R')^2 + (z' + q_z')^2}\ .$$

For corresponding infinitesimal areas on the upper and lower headgroup surfaces we have, respectively:

$$da_h^{(u)} = da_i^{(u)}(R; b + l_h), \qquad da_h^{(l)} = da_i^{(l)}(R; q + l_h). \tag{9}$$

Proceed to the infinitesimal volume of the membrane. This volume occupies a cylindrical sector resulted from rotation of the above-considered strip in plane (R, z) by angle $d\varphi$ around z-axis. Similarly to work [8], introduce the reference frame

$$R_\xi = R + \xi b_R(R)$$
$$z_\xi = z + \xi b_Z(R),$$

where $\left(R_\xi, z_\xi\right)$ are the coordinates of the points on the director with its inner end localized at (R, z); variable ξ takes its values in the interval $(0,1)$ (see Fig.3).

For the upper monolayer, the infinitesimal volume $dV_\xi^{(u)}$ related to $d\xi$, as is clear from Fig.3, can be written as $dV_\xi^{(u)} = R_\xi d\varphi \cdot b_z d\xi \cdot \left|d\vec{r}_\xi\right|$, where $\vec{r}_\xi = \vec{\rho} + \xi\vec{b}$, so that $d\vec{r}_\xi = \left[\left(1 + \xi b_R'\right)\vec{e} + \left(z' + b_z'\right)\vec{k}\right]dR$. Therefore,

$$dV_u = b_z d\varphi dR \int_0^1 d\xi \left(R + \xi b_R\right)\sqrt{\xi^2\left(b_R'^2 + b_z'^2\right) + 2\xi\left(b_R' + z'b_z'\right) + 1 + z'^2}$$
$$= b_z d\varphi dR \int_0^1 d\xi \left(R + \xi b_R\right)\sqrt{\xi^2 A_2 + \xi A_1 + A_0}$$

where $A_0 = 1 + z'^2$, $A_1 = 2\left(b_R' + z'b_z'\right)$, and $A_2 = b_R'^2 + b_z'^2$.

Introduce the designations:

$$I(R)=\int_0^1 d\xi\sqrt{\xi^2A_2+\xi A_1+A_0}=\frac{\left(2A_2+A_1\right)\sqrt{A_2+A_1+A_0}-A_1\sqrt{A_0}}{4A_2}$$
$$+\frac{4A_2A_0-A_1^2}{8A_2^{3/2}}\ln\left|\frac{2A_2+A_1+2\sqrt{A_2(A_2+A_1+A_0)}}{A_1+2\sqrt{A_2A_0}}\right|,$$

$$I_1(R)=\int_0^1 d\xi\cdot\xi\sqrt{\xi^2A_2+\xi A_1+A_0}=\frac{\sqrt{\left(A_2+A_1+A_0\right)^3}-\sqrt{A_0^3}}{3A_2}-\frac{A_1}{2A_2}I(R).$$

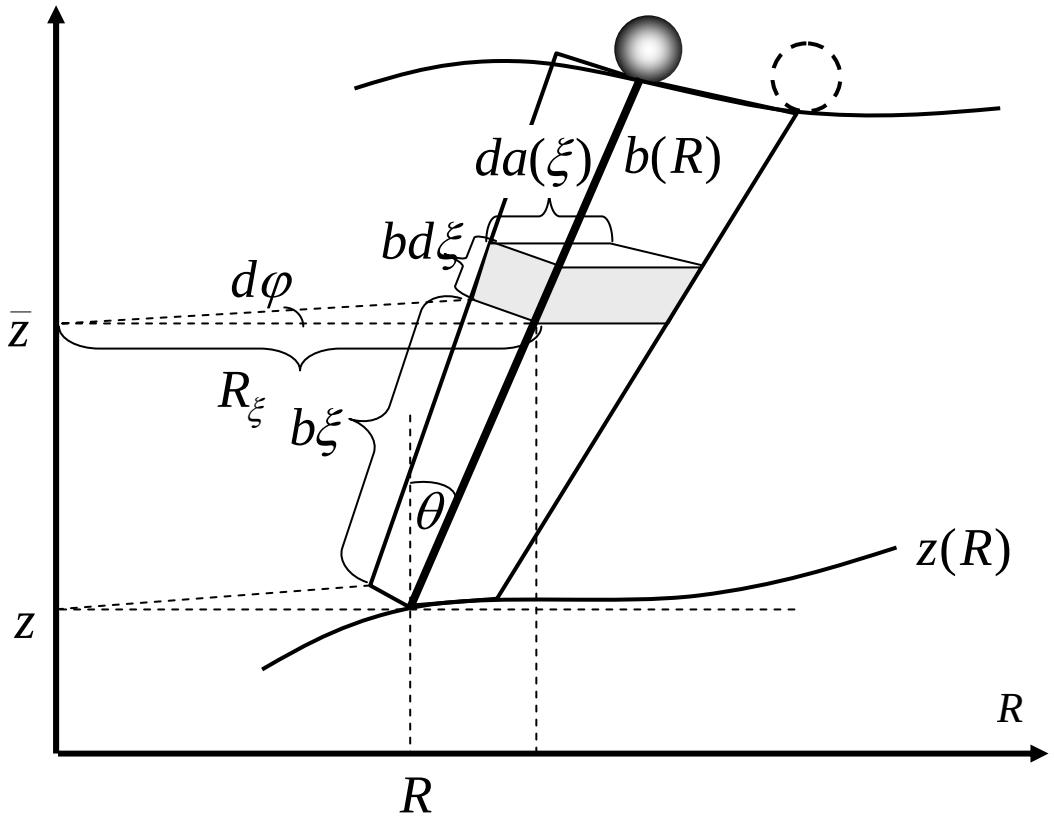


Fig.3.

Then

$$dV_u=b_z d\varphi dR\left[\left(R-b_R\frac{A_1}{2A_2}\right)I(R)+b_R\frac{\sqrt{\left(A_2+A_1+A_0\right)^3}-\sqrt{A_0^3}}{3A_2}\right]=v_u(R)d\varphi dR,$$

where

$$v_u(R)=b_z\left[\left(R-b_R\frac{A_1}{2A_2}\right)I(R)+b_R\frac{\sqrt{\left(A_2+A_1+A_0\right)^3}-\sqrt{A_0^3}}{3A_2}\right] \qquad (10)$$

is the density of the volume with respect to parameter R. Similar expressions for the lower monolayer can be written by the replacements $b \to q, \quad u \to l$.

3.2. Free energy of the bilayer. According to Eq.(6), for the upper monolayer

$$f_u(R) = \gamma a_i(R) + \frac{B}{a_h(R)} + \tau\left[b(R) - l_c\right]^2,$$

where $f(R)$ is the free energy per lipid whose inner tail falls into the interval $(R, R+dR)$. Then the quantity $\frac{f(R)}{\upsilon}$ is the free energy density (as mentioned above, it is supposed that $\upsilon = const$). Therefore, the free energy of this layer reads:

$$F_M^{(u)} = \frac{1}{\upsilon}\int f_u(R)dV_u = \frac{1}{\upsilon}\int_0^{2\pi} d\varphi \int_0^L f_u(R)v_u(R)dR = \frac{2\pi}{\upsilon}\int_0^L f_u(R)v_u(R)dR\,.$$

To specify $f_u(R)$ in this formula, we need to find the expressions for quantities $a_h^{(u)}$ and $a_i^{(u)}$. For this, we write

$$\begin{aligned} a_i^{(u)}(R) &= \int_{\varphi}^{\varphi+\Delta\varphi}\int_{R}^{R+\Delta R} s_u(R)d\varphi dR \approx s_u(R)\Delta\varphi\Delta R \\ \upsilon &= \int_{\varphi}^{\varphi+\Delta\varphi}\int_{R}^{R+\Delta R} v_u(R)d\varphi dR \approx v_u(R)\Delta\varphi\Delta R \end{aligned} \qquad (11)$$

that is approximately valid if $\Delta R << R_{\min}$, where R_{min} is the minimal local curvature radius of the monolayer in the course of membrane deformation. Note that for a spherical micelle or semi-toroidal edge this condition does not hold rigorously. In such cases $R_{\min} \sim b$, being 10÷20 Å for real lipids. At the same time the average inter-lipid distance is also ~10 Å. Then more accurate calculations of the area per lipid, including integration along the membrane midline, should be performed.

If, nevertheless, the approximation (11) is valid, then, obviously,

$$a_i^{(u)}(R) = \frac{s_u(R;b)}{v_u(R)}\upsilon \quad \text{and} \qquad a_h^{(u)}(R) = \frac{s_u(R;b+l_h)}{v_u(R)}\upsilon\,,$$

where $s_u\left(R;b+l_h\right)=\left[R+b_R\left(1+l_h/b\right)\right]\sqrt{\left(1+b_R'\right)^2+\left(z'+b_z'\right)^2}$.

Consequently, the free energy per lipid is

$$f_u(R)=\gamma\upsilon\frac{s_u(R;b)}{v_u(R)}+\frac{B}{\upsilon}\cdot\frac{v_u(R)}{s_u(R;b+l_h)}+\tau\left[b(R)-l_c\right]^2,$$

and the total free energy of the upper monolayer takes the form

$$F_M^{(u)}=\frac{1}{\upsilon}\int f_u(R)dV_u=2\pi\int_0^L\left\{\gamma s_u(R,b)+\frac{B}{\upsilon^2}\cdot\frac{v_u^2(R)}{s_u(R;b+l_h)}+\frac{\tau}{\upsilon}v_u(R)\left[b(R)-l_c\right]^2\right\}dR. \qquad (12)$$

For a planar monolayer the free energy per lipid reads:

$$f_0=\gamma a+\frac{B}{a}+\tau\left(b-l_c\right)^2,\quad ab=\upsilon \qquad (13)$$

(here the index u is omitted in view of full symmetry of a planar bilayer). Also, it is obvious that $dV_u=bRdRd\varphi,\quad da_i^{(u)}=RdRd\varphi$, so that $v_u(R)=bR$ and $s_u(R)=R$. In this case f_0 does not depend on R, and the free energy of a planar monolayer reads:

$$F_M^{(0)}=\frac{2\pi}{\upsilon}f_0\int_0^L bRdR=\pi L^2\frac{b}{\upsilon}f_0=Nf_0 \qquad (14)$$

where N is the given number of lipids. The value b_0 (or a_0) can be found from minimization of f_0 (13) that results in the cubic following equation:

$$\gamma a^3-\left(B-2\tau\upsilon l_c\right)a-2\tau\upsilon^2=0. \qquad (15)$$

Obviously, $L=\sqrt{N\upsilon/\pi b_0}=\sqrt{Na_0/\pi}$.

Introduce dimensionless quantities, with b_0 serving as a length unit:

$$\overline{l_c}=\frac{l_c}{b_0};\quad \overline{B}=\frac{b_0^2}{\gamma\upsilon^2}B;\quad \overline{\tau}=\frac{b_0^3}{\gamma\upsilon}\tau.$$

Then the minimum free energy per lipid in the planar monolayer (at $N = const$) is

$$f_0^{(\min)} = \gamma \frac{\upsilon}{b_0}\left[1 + \bar{B} + \bar{\tau}\left(1 - \bar{l}_c\right)^2\right].$$

Parameters $\bar{\tau}, \bar{B}, \bar{l}_c$, as is seen from Eq.(15) at $a = a_0$, are not independent of each other, precisely,

$$\bar{\tau} = \frac{1}{2} \cdot \frac{1 - \bar{B}}{1 - \bar{l}_c},$$

so that

$$f_0^{(\min)} = \gamma \frac{\upsilon}{b_0}\left[1 + \bar{B} + \frac{1}{2}\left(1 - \bar{B}\right)\left(1 - \bar{l}_c\right)\right]. \qquad (16)$$

Subtracting now (14) with f_0 as (16) from (12) yields the free energy of the upper monolayer deformation $\Delta F_M^{(u)}$. It could be also written in dimensionless quantities, introducing $\bar{R} = R / b_0, \bar{L} = L / b_0, \bar{a} = a / b_0^2$ etc; then $\bar{s}_u(\bar{R}) = s_u(R) / b_0$, $\bar{v}_u(\bar{R}) = v_u(R) / b_0^2$ where $\bar{s}_u(\bar{R})$ are given by Eqs.(8),(10) with all the quantities are supplied with the bar. Finally,

$$\Delta F_M^{(u)} = 2\pi\gamma \int_0^{\bar{L}} d\bar{R} \left\{ \bar{s}_u(\bar{R}; \bar{b}) - \bar{R} + \bar{B}\left[\frac{\bar{v}_u^2(\bar{R})}{\bar{s}_u(\bar{R}; \bar{b} + \bar{l}_h)} - \bar{R} \right] \right.$$
$$\left. + \frac{1}{2} \cdot \frac{1 - \bar{B}}{1 - \bar{l}_c}\left[\left(\bar{b}(\bar{R}) - \bar{l}_c\right)^2 \bar{v}_u(\bar{R}) - \left(1 - \bar{l}_c\right)^2 \bar{R} \right] \right\}$$

Naturally, the expression for the free energy of the lower monolayer can be written with replacing $u \to l, \quad b \to q$.

4. Contribution of the CPP electrostatic potential

The local charge density on the maximum headgroups interaction surface of the upper monolayer can be written as

$$\sigma_u(R) = \frac{Z}{a_h(R)} = \frac{Z}{a_i(R; b + l_h)} = \frac{Z v_u(R)}{s_u(R; b + l_h)\upsilon},$$

where Z is the total headgroup charge, and the other quantities are found above, see (7-10). Then the electrostatic part of the free energy originated from the interaction between the peptide and the upper monolayer reads:

$$F_{\text{int}} = \int_{S_u} \sigma_u \Psi da_h = \int_{S_u} \frac{Z\Psi da_i}{a_i(R; b + l_h)} = \frac{2\pi Z}{\upsilon} \int_0^L \frac{\Psi(R) s_u(R; b)}{s_u(R; b + l_h)} v_u(R) dR, \qquad (13)$$

where $\Psi(\vec{r})$ is the electric potential at point $\vec{r}$.

The real distribution of the electric field created by the peptide near the membrane depends on polarizability and composition of the solvent. The usual tools in allowing for these factors are mean-field theories, like based on the Poisson-Boltzmann equation or its simplest version (Debye screening). At the initial stage of calculations we restrict ourselves with the latter.

While potential $\Psi(\vec{r})$ ensures attraction between the peptide and charged membrane surface, the peptide surface is supposed impermeably rigid. We will model the peptide by a two-axis ellipsoid resulted from rotation of the ellipse with axes $d_{xy} \equiv d_R$ and d_z around z-axis (with the ellipse center located at point z_p and the ellipse axis d_z running along z-axis; for definiteness, $d_z \geq d_R$). Then the ellipsoid surface is composed of points (x, y, z) satisfying the equation

$$\frac{x^2 + y^2}{d_R^2} + \frac{(z - z_p)^2}{d_z^2} = 1.$$

With the relationships $d_z = \varepsilon d$, $d_R = \left(\sqrt{\varepsilon^2 - 1}\right) d$, where $d = \sqrt{d_R^2 + d_z^2}$ is the half–distance between the ellipse foci and $\varepsilon = \frac{d_z}{\sqrt{d_z^2 - d_R^2}}$ is the ellipse elongation parameter, the ellipse surface equation can be re-written as

$$\frac{x^2 + y^2}{\varepsilon^2} + \frac{(z - z_p)^2}{\varepsilon^2 - 1} = d^2.$$

We will assume that there are n_p elementary charges localized at points $\vec{Q}_n = X_n\vec{i} + Y_n\vec{j} + Z_n\vec{k}$ $\left(n = 1,2,..,n_p\right)$ inside the ellipsoid. Then the electric potential at point $\vec{r} = r_x\vec{i} + r_y\vec{j} + r_z\vec{k}$, created by these charges, is $\Psi(\vec{r}) = \sum_n \frac{1}{\left|\vec{r} - \vec{R}_n\right|}\exp\left(-\frac{\left|\vec{r} - \vec{R}_n\right|}{\lambda_D}\right)$ where λ_D is the corresponding Debye length.

To preserve cylindrical symmetry, we suppose that the peptide charges are uniformly situated in the interval $\left\{0,0,z_p + d \geq z \geq z_p - d\right\}$ between the ellipse foci (Fig.4). Then $X_n = Y_n = 0$, and $Z_n = z_p - d + \frac{2d}{n_p - 1}(n-1)$. Therefore,

$$\Psi(\vec{r}) = \Psi(r, r_z) = \sum_{n=1}^{n_p} \frac{\exp\left[-\frac{1}{\lambda_D}\sqrt{r^2 - 2r_z Z_n + Z_n^2}\right]}{\sqrt{r^2 - 2r_z Z_n + Z_n^2}}$$

where $r_z(R) = z(R) + b_z(R)$.

For the membrane surface, obviously

$$\begin{aligned}\vec{r}(R) &= \vec{i}\left[R + b_R(R)\right]\cos\varphi + \vec{j}\left[R + b_R(R)\right]\sin\varphi + \vec{k}\left[z(R) + b_z(R)\right] \\ &= \vec{e}\left[R + b_R(R)\right] + \vec{k}\left[z(R) + b_z(R)\right] \\ r(R) &= \sqrt{R^2 + b^2(R) + z^2(R) + b_z^2 + 2\left[Rb_R(R) + z(R)b_z(R)\right]}.\end{aligned}$$

The proposed way of defining the peptide shape and charge distribution allows us to vary the shape from spherical ($d = 0$) to cylindrical ($\varepsilon = 1$) as well as the distance $d(\varepsilon - 1)$ from the last charge (located in the focus) to the lowest ellipse point $\left\{0,0,z_p - d_z\right\}$. This is helpful in modeling real peptides of different size, shape and charge.

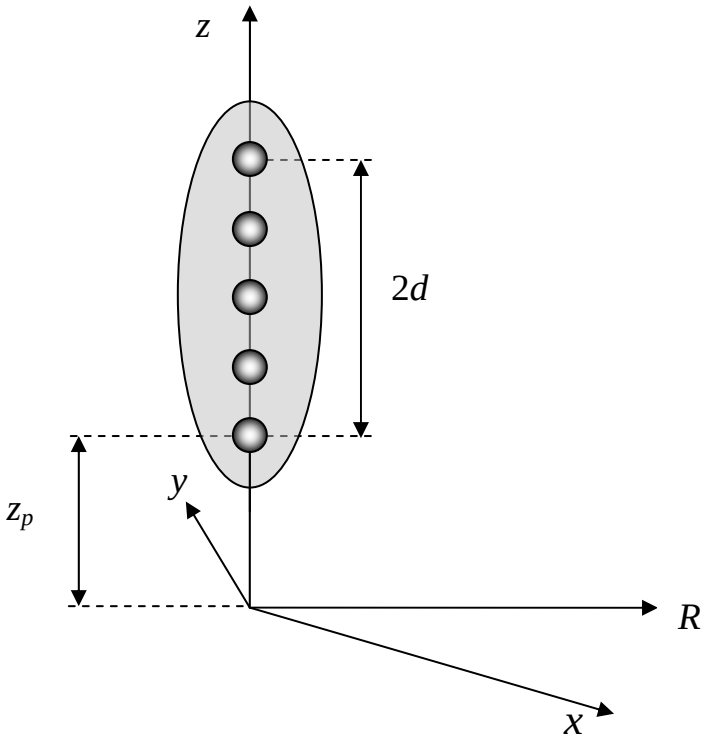


Fig.4.

5. Methods of computation

We now need to obtain the system free energy profile as the peptide approaches the membrane. This task will be done by minimization of the free energy functional $\Delta F_M + F_{\text{int}} = F(z(R), b_R(R), b_z(R), q_R(R), q_z(R))$ with respect to five functions $z(R)$, $b_R(R)$, $b_z(R)$, $q_R(R)$, and $q_z(R)$ for every fixed position of the peptide. Obviously, this cannot be done analytically, and the following standard numerical will be exploited.

The independent variable R is discretized as $R^{(i)}$, $i = 0,1\ldots N$, $R^{(i)} = 0$, $R^{(N)} = L$. The boundary distance L is taken large enough to ensure that all functions at point R_N are the same as those of the unperturbed membrane. This results in the following set of boundary conditions:

$$z(L) = z'(L) = 0, \qquad b_R(L) = b_R'(L) = b_z'(L) = 0, \quad b_z(L) = b_0,$$
$$q_R(L) = q_R'(L) = q_z'(L) = 0, \quad q_z(L) = -b_0,$$

where the quantities with zero subscript are related to the unperturbed membrane. The peptide center is located at point $(0, z_p)$. All the functions characterizing the membrane are properly discretized, too. The *i*-th discrete part of the membrane is characterized by five parameters $z_i, b_{Ri}, b_{zi}, q_{Ri}$, and q_{zi}, with each being varied independently. The total number of variable

parameters is $5N+5$, and functional F is transformed to the function of $5N+5$ variables, $F(z^{(0)}\dots z^{(N)}, b_R^{(0)}\dots b_R^{(N)}, b_z^{(0)}\dots b_z^{(N)}, q_R^{(0)}\dots q_R^{(N)}, q_z^{(0)}\dots q_z^{(N)})$. There are several possible computational approaches to this multiparametric optimization problem. We will consider two of them, most robust and easy to implement.

1. The approach of local variations (also known as unconditional multi-parametric optimization). Each of five parameters are changed by small quantity $\pm\Delta$ at particular point *i* and the new value F_{new} of the functional is computed. If $F_{new} < F$, then the change of the parameter is accepted, if not – rejected. The procedure is applied subsequently to all *N* points until a given convergence criterion is satisfied. The method is simple in implementation but very demanding computationally (although several techniques are known to improve the performance). The boundary conditions are easy to implement in this method.

2. The approach of non-local variations (shape variations). All discretized functions are approximated by finite sets of K cosines

$$z^{(i)} = \sum_{j=0}^{K} \tilde{z}^{(j)} \cos(\omega^{(j)} R^{(i)}), \quad b_R^{(i)} = \sum_{j=0}^{K} \tilde{b}_R^{(j)} \cos(\omega^{(j)} R^{(i)}), \quad b_z^{(i)} = \sum_{j=0}^{K} \tilde{b}_z^{(j)} \cos(\omega^{(j)} R^{(i)}),$$

$$q_R^{(i)} = \sum_{j=0}^{K} \tilde{q}_R^{(j)} \cos(\omega^{(j)} R^{(i)}), \quad q_z^{(i)} = \sum_{j=0}^{K} \tilde{q}_z^{(j)} \cos(\omega^{(j)} R^{(i)}),$$

where ω_i is a set of "harmonics" identical for all parameters; $\tilde{z}^{(j)}, \tilde{b}_R^{(j)}, \tilde{q}_R^{(j)}, \tilde{b}_z^{(j)}, \tilde{q}_z^{(j)}$ are the amplitudes, $\omega^{(0)} = 0$. As a result, F becomes a function of $5N+5$ amplitudes $F(\tilde{z}^{(0)}\dots\tilde{z}^{(K)}, \tilde{b}_R^{(0)}\dots\tilde{b}_R^{(K)}, \tilde{b}_z^{(0)}\dots\tilde{b}_z^{(K)}, \tilde{q}_R^{(0)}\dots\tilde{q}_R^{(K)}, \tilde{q}_z^{(0)}\dots\tilde{q}_z^{(K)})$. The amplitudes are varied to minimize F. The number of harmonics can be quite small. The reason is that the radius of curvature of the membrane is limited – it is larger than b_0 but smaller than L, thus few harmonics can be enough to approximate any possible smooth shape of the membrane. This means that the number of variables is reduced and the computational intensity is substantially lower than for the scheme of local variations. In addition, all derivatives are expressed analytically, and this also results in substantial speed-up of calculations.

The boundary conditions are transformed to

$$\begin{cases} \sum_{j=0}^{K} \tilde{z}^{(j)} \cos(\varpi^{(j)} L) = z^{(0)} \\ \sum_{j=0}^{K} \tilde{z}^{(j)} \omega^{(j)} \sin(\varpi^{(j)} L) = 0 \\ \dots \\ \dots \end{cases}$$

The remaining eight conditions designated by dots can be written by changing $\tilde{z}^{(j)}$ to $\tilde{b}_R^{(j)}, \tilde{q}_R^{(j)}, \tilde{b}_z^{(j)}$ and $\tilde{q}_z^{(j)}$.

It is obvious that the amplitudes can not be varied independently without violating the boundary conditions. At least two amplitudes should be changed consistently at each step. This variational scheme is much more complicated in implementation as compared with that of local variations. The usage of shape variations is justified if the computational burden of local variations scheme becomes intolerable.

Summary

The strategy of analytical and numerical modeling of the interaction of a cell-penetrating peptide with a lipid bilayer membrane is determined and substantiated. As a basis, we choose the extended opposing forces model which we essentially modify in order to allow for considerable membrane geometry changes induced by the peptide. The membrane is considered as a lipid fluid, with lipids length and orientation depending on their positions. Besides, the possibility of deformation of the membrane as a whole due to the bending of the interlayer surface is also taken into consideration. To take account for the disturbing influence of the peptide we propose a simple three-parameter model of the latter, allowing us to vary both peptide geometry (that can be decisive at close contact with the membrane) and its electrostatic field.

The proposed formalism is aimed at determination of the system free energy profile and the character of membrane deformation with the peptide approaching. This will be done by minimization of the constructed free energy functional of five independent functions defining the equilibrium membrane configuration in a convenient and visual way. Several minimization procedures will be tested to find the optimal one.

We suppose that the free energy profile to be found will allow us to determine the optimal characteristics of the peptide in regard to its transduction abilities. This will also make it possible to construct a diffusion kinetic model of the transduction process and to evaluate its kinetic parameters.

References

1. W. Helfrich. Elastic properties of lipid bilayers: Theory and possible experiments. *Z. Naturforsch.* **28C**, 693-703 (1973).
2. J.N. Israelachvili, D.J. Mitchell, B.W. Ninham Theory of self-assembly of hydrocarbon amphiphiles into micelles and bilayers. *J. Chem. Soc. Faraday Trans.* **72**, 1525-1568 (1976).
3. J.N. Israelachvili. *Intermolecular and Surface Forces*, 2nd ed. Academic Press, 1992.
4. J.N. Israelachvili, I. Ladyzinski. The physico-chemical basis of self-assembling structures. In: *Forces, Growth and Form in Soft Condensed Matter: At the Interface between Physics and Biology* (A.T.Skjeltorp, A.V. Belushkin, eds.). Kluwer Academic, 2004, pp.1-28.
5. S. May, Y. Bohbot, A. Ben-Shaul. Molecular theory of bending elasticity and branching of cylindrical micelles. *J. Phys. Chem. B* **101**, 8648-8657 (1997).
6. S. May, A. Ben-Shaul. Molecular theory of lipid-protein interaction and the $L_\alpha - H_{\square}$ transition. *Biophys. J.* **76**, 751-767 (1999).
7. A. Zemel, D.R. Fattal, A. Ben-Shaul. Energetics and self-assembly of amphipathic peptide pores in lipid membranes. *Biophys. J.* **84**, 2242-2255 (2003).
8. S. May. A molecular model for the line tension of lipid membranes. *Eur. Phys. J. E* **3**, 37-44 (2000).
9. A. Ben-Shaul. Molecular theory of chain packing, elasticity and lipid-protein interactions in lipid bilayers. In: *Handbook of Biological Physics*, V.1: *Structure and Dynamics of Membranes* (R. Lipowski, E. Sackmann, eds.). Elsevier, 1995, Ch.7, pp.359-401.
10. S. May, D. Harries, A. Ben-Shaul. Lipid demixing and protein-protein interactions in the adsorption of charged proteins on mixing membranes. *Biophys. J.* **79**, 1747-1760 (2000).
11. D. Harries, A. Ben-Shaul, I. Szleifer. Enveloping of charged proteins by lipid bilayers. *J. Phys. Chem. B* **108**, 1491-1496 (2004).
12. D. Andelmann. Electrostatic Properties of Membranes: The Poisson-Boltzmann Theory. In: *Handbook of Biological Physics*, V.1: *Structure and Dynamics of Membranes* (R. Lipowski, E. Sackmann, eds.). Elsevier, 1995.
13. S. May, A. Ben-Shaul. Spontaneous curvature and thermodynamic stability of mixed amphiphilic layers. *J. Chem. Phys.* **103**, 3839-3848 (1995).
14. C. Tanford. The Hydrophobic Effect. Wiley, 1980.

EOARD Partner Project 211 – Deliverable 3

1. Introduction

Previously (see Deliverables 1,2) we have analyzed the known theoretical approaches to studying the dependence of the shape of bilayer phospholipid membranes on various disturbing factors in order to choose the most promising method of the theoretical investigation of the CPP transduction process. As a result, we have chosen a version of the phenomenological opposing forces model (OFM) [1-3] having much in common with the model of oriented liquid. The main advantage of this model consists in its comparative simplicity and supposed ability to treat considerable deformations of the bilayer. True, in recent papers (see e.g. [4]) the OFM was used mainly for description of small perturbations of a given membrane geometry. Needless to say, studying large (even if local) changes of the membrane shape, one implies that the model exploited is able to describe the initial unperturbed membrane configuration. At the same time such a model must be instructive enough at describing the emergence of rather complex membrane structures of nontrivial geometry, up to formation of pores/channels for CPPs. That is why in Delivery 2 we have presented a special mathematical formalism, aimed directly at such problems, within which the bilayer lipid membrane configurations are described with five independent functions determining both the average characteristics of orientation and deformation of the lipids in each monolayer as well as bends of the membrane as a whole.

The present stage of the project is devoted to creating the programming methods of realization of the model developed in order to calculate and simulate the membrane configurations at various values of the model parameters, for both the unperturbed membrane and that interacting with the CPP. At this stage the problem consists in numerical minimization of the system free energy functional, with the membrane configuration given by the five mentioned functions resulted from the minimization procedure.

To this end, we go two different ways. The first is to apply the known Euler-Lagrange formalism. This leads to the necessity of numerically solving the set of ten nonlinear first-order ODE for the five mentioned functions and their first derivatives. The second consists in direct variations of the lipid arrangement in the membrane.

2. Minimization of the membrane free energy functional by the Euler-Lagrange method

According to the results presented in Deliverable 2, within the OFM model the free energy of a membrane of arbitrary cylindric-symmetrical configuration reads

$$F_M = 2\pi\gamma \int_0^L \mathrm{L}\ \left(R; z, z', b_R, b_z, b'_R, b'_z, q_R, q_z, q'_R, q'\right) dR \tag{1}$$

where the Lagrange function of the membrane

$$\begin{aligned} \mathrm{L}\ \left(R, z, z', b_R, b'_R, b_z, b'_z, q_R, q'_R, q_z, q'_z\right) = \mathrm{L}_{\mathrm{u}}\left(R, z, z', b_R, b'_R, b_z, b'_z\right) \\ + \mathrm{L}_{\mathrm{l}}\left(R, z, z', q_R, q'_R, q_z, q'_z\right) \end{aligned} \tag{2}$$

consists of the two terms related to the "upper" (u) and "lower" (l) monolayer. Functions $b_R(R), b_z(R)$ represent the corresponding components of vector $\vec{b}$, determining the length and orientation of a lipid at point R in the upper monolayer (see Fig.2 of Delivery 2), and $b'_R(R), b'_z(R)$ are their derivatives with respect to R; $q_R(R), q_z(R)$ and $q'_R(R), q'(R)$ are the analogues for the lower monolayer. Lastly, functions $z(R), z'(R)$ determine the shape of the midline between the monolayers.

Under cylindric symmetry, for e.g. the upper monolayer we have obtained:

$$\mathrm{L}_{\mathrm{u}}\left(R, z, z', b_R, b'_R, b_z, b'_z\right) = s_u\left(R; b\right) + B\frac{v_u^2(R)}{s_u\left(R; b + l_h\right)} + \frac{1}{2}\cdot\frac{1-B}{1-l_c}\left(b(R) - l_c\right)^2 v_u(R, b) \tag{3}$$

(see Deliverable 2 for the conventional constant designations), where

$$s_u(R, \vec{b}) = \left(R + b_R\right)\sqrt{\left(1 + b'_R\right)^2 + \left(z' + b'_z\right)^2} \tag{4}$$

with $s_u(R, \vec{b})dR$ being the infinitesimal hydrophobic surface corresponding to the cylindric-symmetrical infinitesimal volume $v_u(R)dR$, where

$$v_u(R) = b_z \int_0^1 d\xi \left(R + \xi b_R\right)\sqrt{\left(1 + \xi b'_R\right)^2 + \left(z' + \xi b'_z\right)^2}. \tag{5}$$

The corresponding analogues for the lower monolayer can be written by the replacement of the components of vector $\vec{b}(R)$ and their derivatives by those of vector $\vec{q}(R)$.

The Euler-Lagrange equations for the sought functions $z(R), b_R(R), b_z(R), q_R(R), q_z(R)$ read:

$$\frac{\mathrm{d}}{\mathrm{d}R}\frac{\partial \mathrm{L}}{\partial X'} = \frac{\partial \mathrm{L}}{\partial X}, \qquad X = z, b_R, b_z, q_R, q_z. \tag{6}$$

Taking into account that $\frac{\partial \mathrm{L}}{\partial z'} = 0$ and also that

$$\frac{\mathrm{d}}{\mathrm{d}R} = \frac{\partial}{\partial R} + z'' \frac{\partial}{\partial z'} + z' \frac{\partial}{\partial z} + b_R'' \frac{\partial}{\partial b_R'} + b_z'' \frac{\partial}{\partial b_z'} + b_R' \frac{\partial}{\partial b_R} + b_z' \frac{\partial}{\partial b_z} + q_R'' \frac{\partial}{\partial q_R'} + q_z'' \frac{\partial}{\partial q_z'} + q_R' \frac{\partial}{\partial q_R} + q_z' \frac{\partial}{\partial q_z'}$$

we arrive at the set of Euler-Lagrange equations in the form

$$\begin{aligned}
&\frac{dy_1}{dx} = y_2 \\
&\frac{dy_2}{dx} L_{22} + \frac{dy_4}{dx} L_{24} + \frac{dy_6}{dx} L_{26} + \frac{dy_8}{dx} L_{28} + \frac{dy_{10}}{dx} L_{210} = B_2 \\
&\frac{dy_3}{dx} = y_4 \\
&\frac{dy_2}{dx} L_{24} + \frac{dy_4}{dx} L_{44} + \frac{dy_6}{dx} L_{46} = B_4 \\
&\frac{dy_5}{dx} = y_6 \\
&\frac{dy_2}{dx} L_{26} + \frac{dy_4}{dx} L_{46} + \frac{dy_6}{dx} L_{66} = B_6 \\
&\frac{dy_7}{dx} = y_8 \\
&\frac{dy_2}{dx} L_{28} + \frac{dy_8}{dx} L_{88} + \frac{dy_{10}}{dx} L_{810} = B_8 \\
&\frac{dy_9}{dx} = y_{10} \\
&\frac{dy_2}{dx} L_{210} + \frac{dy_8}{dx} L_{810} + \frac{dy_{10}}{dx} L_{1010} = B_{10}
\end{aligned} \tag{7}$$

where

$$\begin{aligned}
&x = R; \quad y_i = y_i(x) \qquad z = y_1 \quad z' = y_2 \\
&b_R = y_3 \quad b_R' = y_4 \qquad q_R = y_7 \quad q_R' = y_8 \\
&b_z = y_5 \quad b_z' = y_6 \qquad q_z = y_9 \quad q_z' = y_{10}
\end{aligned} \tag{8}$$

and also

$$\frac{\partial \mathrm{L}}{\partial y_i} = \mathrm{L}_i, \quad \frac{\partial^2 \mathrm{L}}{\partial R \partial y_i} = \mathrm{L}_{Ri}, \quad \frac{\partial^2 \mathrm{L}}{\partial y_i \partial y_k} = \mathrm{L}_{ik}. \tag{9}$$

These quantities expressed in terms of the components of vectors $\vec{b}(R)$ or/and $\vec{q}(R)$ are given in Appendix. Besides,

$$
\begin{aligned}
B_2 &= -L_{02} - y_4 L_{23} - y_6 L_{25} - y_8 L_{27} - y_{10} L_{29} \\
B_4 &= L_3 - L_{04} - y_4 L_{34} - y_6 L_{45} \\
B_6 &= L_5 - L_{06} - y_4 L_{36} - y_6 L_{56} \\
B_8 &= L_7 - L_{08} - y_8 L_{78} - y_{10} L_{89} \\
B_{10} &= L_9 - L_{010} - y_8 L_{710} - y_{10} L_{910}.
\end{aligned}
\tag{10}
$$

Inserting (10) into (4) yields

$$
\frac{dy_i}{dx} = f_i(x, y_1, \ldots y_{10}), \qquad i = 1, \ldots 10 \tag{11}
$$

where

$$
\begin{aligned}
f_1(y_2) &= y_2 \\
f_2(x, y_1, \ldots y_{10}) &= \frac{1}{Q}\left[B_2 - \frac{B_4\left(L_{24}L_{66} - L_{26}L_{46}\right) + B_6\left(L_{26}L_{44} - L_{24}L_{46}\right)}{L_{44}L_{66} - L_{46}^2} \right. \\
&\qquad \left. - \frac{B_8\left(L_{28}L_{1010} - L_{210}L_{810}\right) + B_{10}\left(L_{210}L_{88} - L_{28}L_{810}\right)}{L_{88}L_{1010} - L_{810}^2} \right] \\
f_3(y_4) &= y_4 \\
f_4(x, y_1, \ldots y_{10}) &= \frac{B_4 L_{66} - B_6 L_{46} - f_2(x, y_1, \ldots y_{10})\left(L_{24}L_{66} - L_{26}L_{46}\right)}{L_{44}L_{66} - L_{46}^2} \\
f_5(y_6) &= y_6 \\
f_6(x, y_1, \ldots y_{10}) &= \frac{B_6 L_{44} - B_4 L_{46} - f_2(x, y_1, \ldots y_{10})\left(L_{44}L_{26} - L_{24}L_{46}\right)}{L_{44}L_{66} - L_{46}^2} \\
f_7(y_8) &= y_8 \\
f_8(x, y_1, \ldots y_{10}) &= \frac{B_8 L_{1010} - B_{10} L_{810} - f_2(x, y_1, \ldots y_{10})\left(L_{28}L_{1010} - L_{210}L_{810}\right)}{L_{88}L_{1010} - L_{810}^2} \\
f_9(y_{10}) &= y_{10} \\
f_{10}(x, y_1, \ldots y_{10}) &= \frac{B_{10} L_{88} - B_8 L_{810} - f_2(x, y_1, \ldots y_{10})\left(L_{88}L_{210} - L_{28}L_{810}\right)}{L_{88}L_{1010} - L_{810}^2}
\end{aligned}
\tag{12}
$$

and

$$Q = L_{22} - \frac{L_{24}^2 L_{66} + L_{26}^2 L_{44} - 2L_{24}L_{26}L_{46}}{L_{44}L_{66} - L_{46}^2} - \frac{L_{28}^2 L_{1010} + L_{210}^2 L_{88} - 2L_{28}L_{210}L_{810}}{L_{88}L_{1010} - L_{810}^2}. \quad (13)$$

To numerically solve set (11) we have developed a program using the Runge-Kutta method with adaptive step size. In the vicinity of the zero point the integrals were computed by the Gauss method with prescribed accuracy.

We have obtained preliminary results showing that there exists some domain in the system parameters space where the solution can be found in a narrow interval of variable R; however, extending this interval leads to solution instability (see Fig.1 as an example). The lattr is caused by a series of singularities like zero coefficients at the highest derivatives, etc. At present, the numerical scheme is under modernization aimed at removing such drawbacks.

3. Minimization of the free energy functional by the method of local variations

At present, the program realizing the method of local variations (see Delivery 2) is at its testing-debugging stage. This program allows us to obtain the optimal configuration of the free membrane and monitor the changes of different components of the system energy. The program is implemented in two modifications which use direct variations or those in the Fourier space of the initial functions components.

These program versions are realized in the Object Pascal language in Delphi 5 Enterprise integrated development environment. The elementary surfaces are calculated with the linearized expressions (see Deliverable 2). The elementary volumes are calculated by numerical integration of the corresponding expressions with the help of the standard method of trapezoids.

In the first version we use the method of local variations under the integral condition of volume constancy whereas in the second – the method of shape variations under the same condition. Both versions use the standard method of unconditional successive multidimensional optimization in the presence of local steric restrictions. The exemplary screenshots are given below (see Figs.2-4).

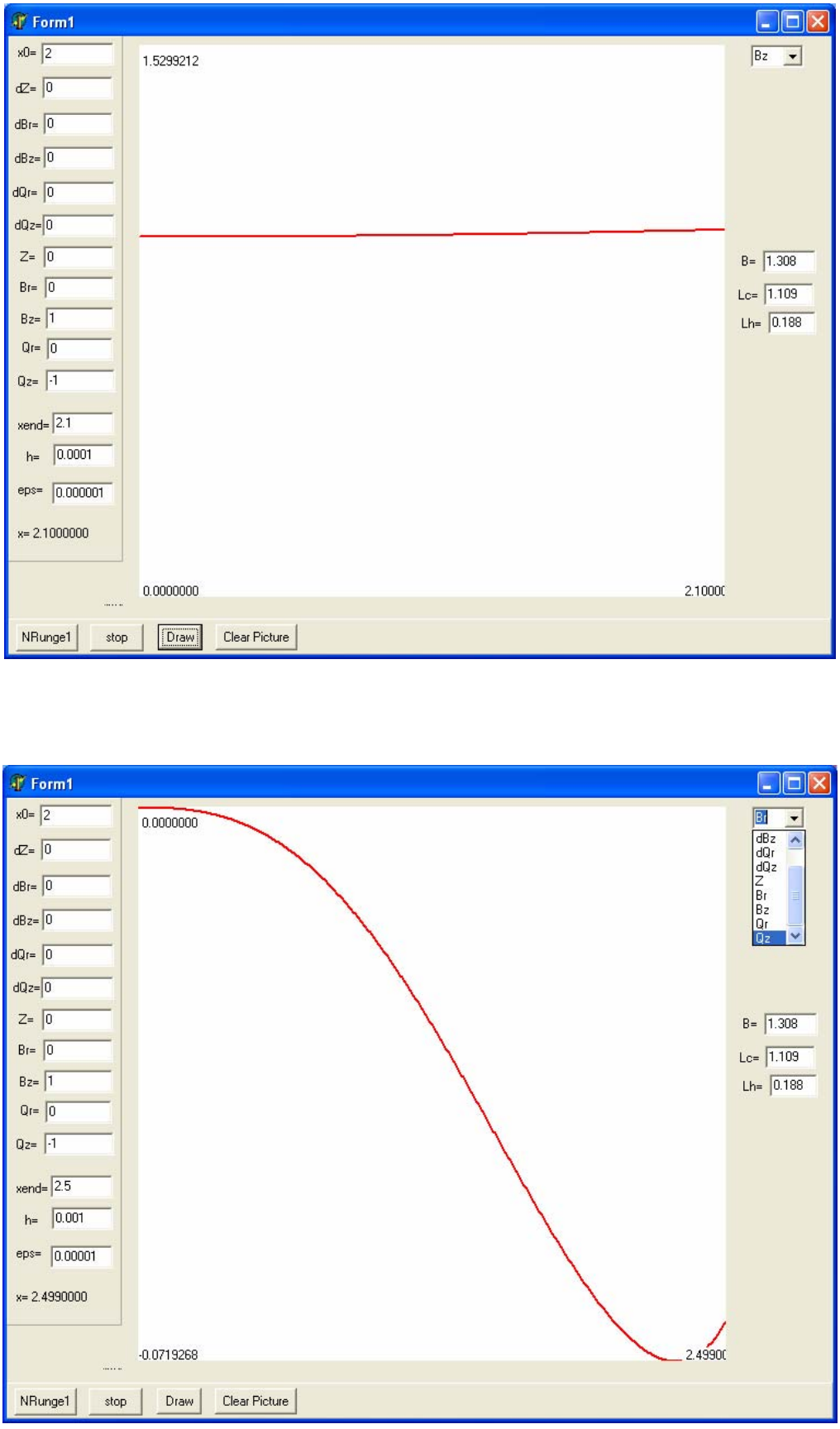


Fig.1. Modelling of a planar bilayer (note a difference in xend in the upper and lower screenshots).

Fig.2. The beginning of the optimization procedure (method of local variations).

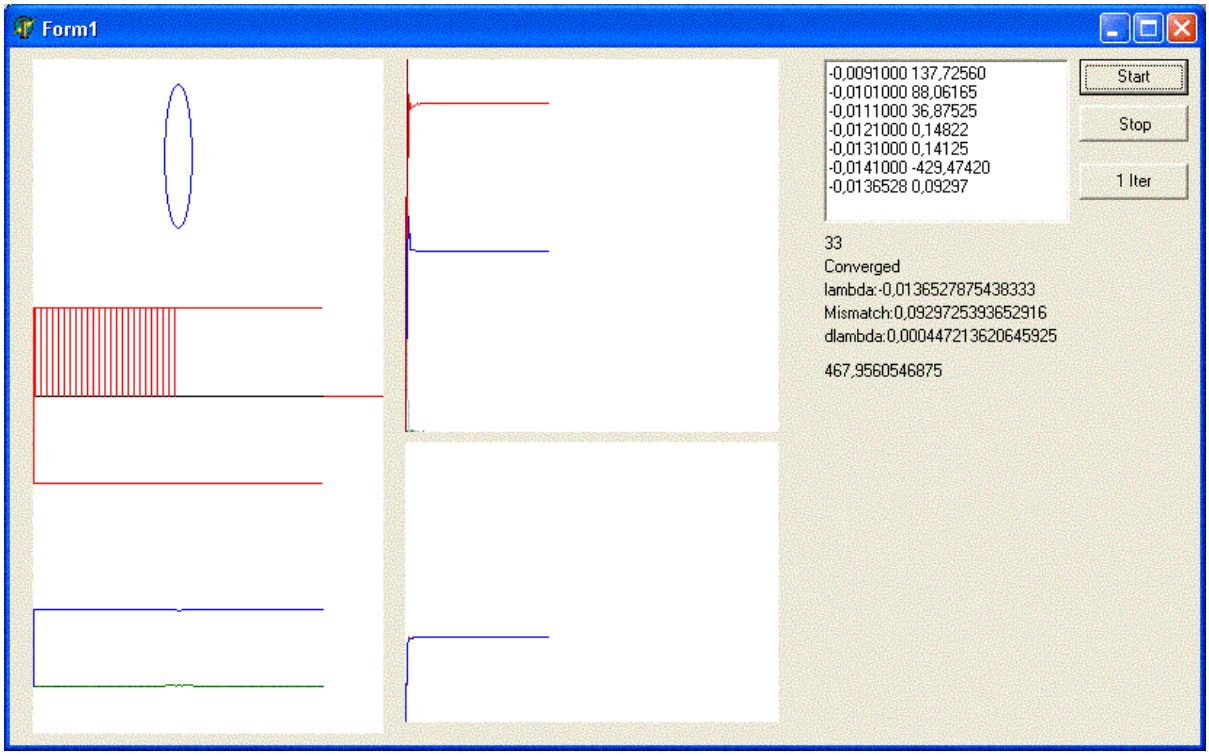


Fig.3. The final result (equilibrium planar bilayer).

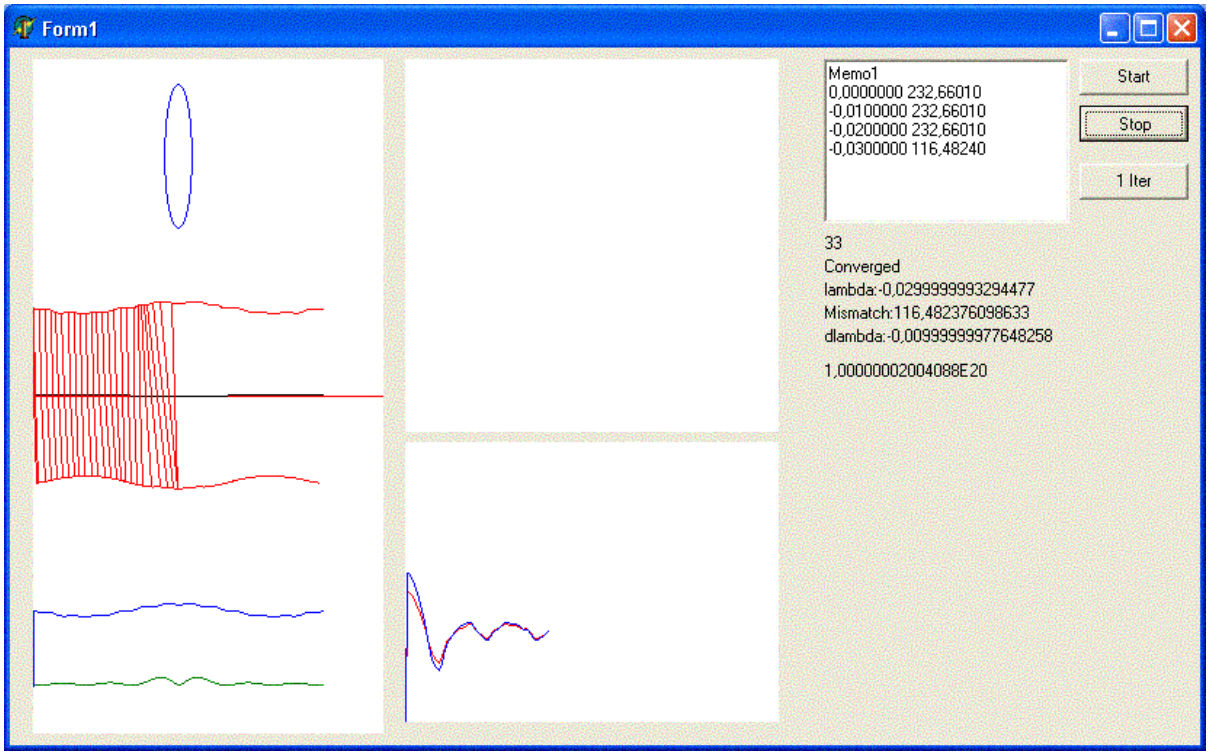


Fig.4. The course of optimization in the Fourier space.

The advantage of the method of shape variation consists in the guaranteed absence of punctual disruptions of derivatives which destabilize the computational scheme. Yet, each iteration requires full recalculation of the system geometry and energy parameters of the whole membrane, essentially impeding the program execution.

In both methods the principle of steric overlap exclusion for lipid headgroups situated in two neighbouring discretization points holds. To test the presence of such overlap, the following relationship for each pair of neighbouring discretization points i and $i+1$ is verified:

$$\begin{cases} d_i > l_i \\ d_{i+1} > l_{i+1} \end{cases} \qquad (14)$$

where d is the distance from the membrane midline to the point of crossing of directors of corresponding lipids, $l = b + l_h$ is the whole length of the lipid (see Fig.5). If this relationship does not hold, then the corresponding variation causing the steric overlap is rejected.

The volume constancy condition is introduced with the help of the standard Lagrange method of undetermined multipliers. The following term is inserted into the free energy functional:

$$f_V = \lambda \int \left(v(R) - v_0(R) \right) dR$$

where λ is the Lagrange multiplier and $v_0(R)dR$ is the infinitesimal volume of the planar unperturbed membrane. The optimization problem is solved for a series of values λ unless the value of пока значение f_V becomes sufficiently close to zero. The search of the corresponding value of λ is basically a standard task of finding the zero of a function of one variable and is performed by bipartitioning.

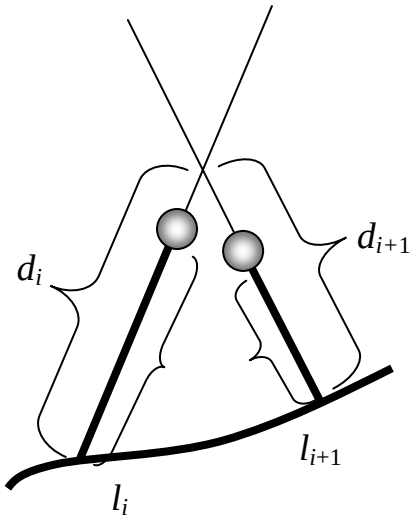


Fig.5. To condition (14).

4. Conclusions

The defined problem of optimization of the membrane shape in the presence of the CPP has turned out to be, methodically and technically, far more complex that it could be expected from the existing literature. Although the chossen approach, based on the extended OFM, in principle allows one to describe complex membrane structures of any shape, yet it has been previously exploited for calculations of mainly small deviations from the planar shape of the bilayer. In those cases the corresponding Euler-Lagrange equations can be linearized and the equilibrium shape of the membrane can be found comparatively easily. The general task of large nonlinear deformations is intrinsically complicated and specific for the following reasons:

(i) the Euler-Lagrange equations are very complex in the general case;

(ii) the solution corresponding to the planar unperturbed membrane represents a singularity of these equations. Consequently, this requires specific regularization procedures;

(iii) some terms of these equations contain integrals with their analytical expressions having critical points in the parameters range of interest. This forces to employ numerical computing of the integrals and leads to far lower computational intensity;

(iv) finally, the functions under optimization are fast-varying and therefore strongly toughen the requirements to numerical algorithms.

An alternative method of optimization – method of local variations – has also turned out quite complex and specific in realization. The corresponding complexities are connected with bifurcational dependences of the optimal solution on parameters values and also with the presence of multiple local minima, nonlocal character of the volume constancy condition as well as with the instability of the standard computational scheme with respect to punctual "defects". In view of this, we have decided to develop both approaches in parallel since none of them is noticeably preferable.

The results obtained are of preliminary character and indicate the necessity of further development of the optimization procedure for the free energy functional in order to construct a stable scheme, yielding first of all the unperturbed/planar membrane shape. Then the numerical experiment on the (bilayer membrane + CPP) system will go comparatively smoothly. The development of sseveral computational approaches seems quite reasonable at the present stage.

The need of some further modernization of the extended OFM model in order to obtain stable solutions for the unperturbed membrane looks also very likely. This will be cleared up in the course of further application of the above-described versions of the optimization schemes. Probably, more powerful computational means than presently used standard PCs will be needed.

Appendix

Here we decipher the designations used above, reflecting the complexity of the problem. Recall that the superscripts of monolayers (supposing identical expressions for both under the replacement $b \to q$) and the bars denoting renormalized dimensionless quantities are omitted. Besides, we also omit the terms of the Lagrangian that contain R only (as full derivatives with respect to R, inessential for minimization) and denote

$$s\left(R;b\right) \equiv s, \quad s\left(R;b+l_h\right) \equiv \tilde{s}.$$

The first derivatives of L read

$$\frac{\partial \mathrm{L}}{\partial \eta} = \frac{\partial s}{\partial \eta} + B\frac{v}{\tilde{s}^2}\left(2\tilde{s}\frac{\partial v}{\partial \eta} - v\frac{\partial \tilde{s}}{\partial \eta}\right) + \frac{\mu}{2}\left(b - l_c\right)\left[2\frac{\partial b}{\partial \eta}v + \left(b - l_c\right)\frac{\partial v}{\partial \eta}\right],$$
$$\eta = z, b_R, b_z, z', b'_R, b'_z.$$

Here $\mu \equiv \dfrac{1-B}{1-l_c}$. In terms of the problem variables

$$s = \left(R + b_R\right)\sqrt{\left(1 + b'_R\right)^2 + \left(z' + b'_z\right)^2} \equiv \left(R + b_R\right)K_1$$
$$\tilde{s} = \left(R + b_R\right)K_1$$
$$K_1 = \sqrt{\left(1 + b'_R\right)^2 + \left(z' + b'_z\right)^2}$$
$$\tilde{K}_1 = \sqrt{\left(1 + \xi b'_R\right)^2 + \left(z' + \xi b'_z\right)^2}$$
$$v = b_z\int_0^1 d\xi\left(R + \xi b_R\right)\sqrt{\left(1 + \xi b'_R\right)^2 + \left(z' + \xi b'_z\right)^2}$$
$$\equiv b_z\int_0^1 d\xi\left(R + \xi b_R\right)\sqrt{\xi^2 A_2 + \xi A_1 + A_0}$$
$$\equiv b_z\left(RI + b_R I_1\right)$$
$$I = \int_0^1 d\xi\sqrt{\xi^2 A_2 + \xi A_1 + A_0} = \int_0^1 \tilde{K}_1 d\xi$$
$$I_1 = \int_0^1 \xi d\xi\sqrt{\xi^2 A_2 + \xi A_1 + A_0} = \int_0^1 \tilde{K}_1 \xi d\xi$$
$$A_0 = 1 + z'^2; \quad A_1 = 2\left(b'_R + z' b'_z\right); \quad A_2 = b'^2_R + b'^2_z.$$

Introduce aslo

$$I_n = \int_0^1 \frac{\xi^{n-2}d\xi}{\tilde{K}_1}$$
$$J_n = \int_0^1 \frac{\xi^{n-2}d\xi}{\tilde{K}_1^3}$$

for $n \geq 2$, and

$$P \equiv R + b_R\left(1 + \frac{l_h}{b}\right) = \frac{\tilde{s}}{K_1}$$
$$R_1 \equiv RI + b_R I_1 = \frac{v}{b_z}$$
$$\Lambda \equiv 1 + l_h\frac{b_z^2}{b^3}$$
$$\Lambda_1 \equiv l_h\frac{b_z b_R}{b^3}.$$

Besides,

$$t_1 \equiv Rz' I_2 + \left(Rb'_z + b_R z'\right) I_3 + b_R b'_z I_4$$
$$t_2 \equiv R I_3 + \left(Rb'_R + b_R\right) I_4 + b_R b'_R I_5$$
$$t_3 \equiv Rz' I_3 + \left(Rb'_z + b_R z'\right) I_4 + b_R b'_z I_5,$$

and of course $b = \sqrt{b_R^2 + b_z^2}$.

Below we list the derivatives entering Eqs.(7).

$$\frac{\partial^2 \mathrm{L}}{\partial {b'_R}^2} = \frac{R + b_R}{K_1}\left[1 - \frac{\left(1 + b'_R\right)^2}{K_1^2}\right] + 2Bb_z^2 \frac{R_1}{PK_1}\frac{\partial t_2}{\partial b'_R} + b_z \frac{\mu}{2}\left(b - l_c\right)^2 \frac{\partial t_2}{\partial b'_R}$$
$$+ 2Bb_z^2 \frac{t_2}{PK_1}\left(t_2 - R_1 \frac{1 + b'_R}{K_1^2}\right) - Bb_z^2 \frac{R_1^2}{PK_1^3} - Bb_z^2 \frac{R_1}{PK_1^3}\left(1 + b'_R\right)\left(2t_2 - 3R_1 \frac{1 + b'_R}{K_1^2}\right)$$

$$\frac{\partial^2 \mathrm{L}}{\partial b_z \partial b'_R} = -\frac{R + b_R}{K_1^3}\left(1 + b'_R\right)\left(z' + b'_z\right) + 2Bb_z^2 \frac{R_1}{PK_1}\frac{\partial t_3}{\partial b'_R} + b_z \frac{\mu}{2}\left(b - l_c\right)^2 \frac{\partial t_3}{\partial b'_R}$$
$$+ 2Bb_z^2 \frac{t_2}{PK_1}\left(t_3 - R_1 \frac{z' + b'_z}{K_1^2}\right) - Bb_z^2 \frac{R_1}{PK_1^3}\left(1 + b'_R\right)\left(2t_3 - 3R_1 \frac{z' + b'_z}{K_1^2}\right)$$

$$\frac{\partial \mathrm{L}}{\partial b_R} = K_1 + 2Bb_z^2 \frac{R_1}{PK_1} I_1 + b_z \frac{\mu}{2}\left(b - l_c\right)^2 I_1 - Bb_z^2 \frac{R_1^2}{P^2 K_1}\Lambda + R_1 \frac{b_z b_R}{b}\mu\left(b - l_c\right)$$

$$\frac{\partial^2 \mathrm{L}}{\partial b_R \partial b'_R} = \frac{1 + b'_R}{K_1} + 2Bb_z^2 \frac{R_1}{PK_1}\frac{\partial t_2}{\partial b_R} + 2Bb_z^2 \frac{t_2}{PK_1}\left(I_1 - R_1 \frac{\Lambda}{P}\right) + b_z \frac{\mu}{2}\left(b - l_c\right)^2 \frac{\partial t_2}{\partial b_R}$$
$$+ \frac{b_z b_R}{b}\mu\left(b - l_c\right) t_2 - Bb_z^2 \frac{R_1}{PK_1^3}\left(1 + b'_R\right)\left(2I_1 - R_1 \frac{\Lambda}{P}\right)$$

$$\frac{\partial^2 \mathrm{L}}{\partial b_z \partial b'_R} = t_2\left[4Bb_z \frac{R_1}{PK_1} + \frac{\mu}{2}\left(b - l_c\right)^2 + \mu\left(b - l_c\right)\frac{b_z^2}{b} + 2Bb_z^2 \frac{R_1 \Lambda_1}{P^2 K_1}\right]$$
$$- Bb_z \frac{R_1^2}{PK_1^3}\left(1 + b'_R\right)\left(2 + b_z \frac{\Lambda_1}{P}\right)$$

$$\frac{\partial^2 \mathrm{L}}{\partial R \partial b'_R} = \frac{1 + b'_R}{K_1} + b_z \frac{\partial t_2}{\partial R}\left[2Bb_z \frac{R_1}{PK_1} + \frac{\mu}{2}\left(b - l_c\right)^2\right] + 2Bb_z^2 \frac{t_2}{PK_1}\left(I - \frac{R_1}{P}\right)$$
$$- Bb_z^2 \frac{R_1}{PK_1^3}\left(1 + b'_R\right)\left(2I - \frac{R_1}{P}\right)$$

$$\frac{\partial^2 \mathrm{L}}{\partial b_z'^2} = \frac{R + b_R}{K_1}\left[1 - \frac{\left(z' + b_z'\right)^2}{K_1^2}\right] + 2Bb_z^2 \frac{R_1}{PK_1}\frac{\partial t_3}{\partial b_z'} + b_z \frac{\mu}{2}\left(b - l_c\right)^2 \frac{\partial t_3}{\partial b_z'}$$
$$+ 2Bb_z^2 \frac{t_3}{PK_1}\left(t_3 - R_1 \frac{z' + b_z'}{K_1^2}\right)$$
$$- Bb_z^2 \frac{R_1}{PK_1^3}\left[2\left(z' + b_z'\right)t_3 + R_1 - 3R_1 \frac{\left(z' + b_z'\right)^2}{K_1^2}\right]$$

$$\frac{\partial \mathrm{L}}{\partial b_z} = Bb_z \frac{R_1^2}{PK_1}\left(2 + b_z \frac{\Lambda_1}{P}\right) + R_1\left[\frac{\mu}{2}\left(b - l_c\right)^2 + \mu\left(b - l_c\right)\frac{b_z^2}{b}\right]$$

$$\frac{\partial^2 \mathrm{L}}{\partial b_R \partial b_z'} = \frac{z' + b_z'}{K_1} + b_z \frac{\partial t_3}{\partial b_R}\left[2Bb_z \frac{R_1}{PK_1} + \frac{\mu}{2}\left(b - l_c\right)^2\right]$$
$$+ b_z t_3\left[2B\frac{b_z}{PK_1}\left(I_1 - R_1 \frac{\Lambda}{P}\right) + \mu\left(b - l_c\right)\frac{b_R}{b}\right]$$
$$- Bb_z^2 \frac{R_1}{PK_1^3}\left(z' + b_z'\right)\left(2I_1 - R_1 \frac{\Lambda}{P}\right)$$

$$\frac{\partial^2 \mathrm{L}}{\partial b_z \partial b_z'} = Bb_z \frac{R_1}{PK_1}\left(2 + b_z \frac{\Lambda_1}{P}\right)\left(2t_3 - R_1 \frac{z' + b_z'}{K_1^2}\right)$$
$$+ t_3\left[\frac{\mu}{2}\left(b - l_c\right)^2 + \mu\left(b - l_c\right)\frac{b_z^2}{b}\right]$$

$$\frac{\partial^2 \mathrm{L}}{\partial R \partial b_z'} = \frac{z' + b_z'}{K_1} + b_z \frac{\partial t_3}{\partial R}\left[2Bb_z \frac{R_1}{PK_1} + \frac{\mu}{2}\left(b - l_c\right)^2\right]$$
$$+ 2Bb_z^2 \frac{t_3}{PK_1}\left(I - \frac{R_1}{P}\right) - Bb_z^2 \frac{R_1}{PK_1^3}\left(z' + b_z'\right)\left(2I - \frac{R_1}{P}\right).$$

$$\frac{\partial \mathrm{L}}{\partial R} = K_1 + Bb_z^2 \frac{R_1}{PK_1}\left(2I - \frac{R_1}{P}\right) + \frac{\mu}{2}\left(b - l_c\right)^2 b_z I$$

$$\frac{\partial^2 \mathrm{L}}{\partial z' \partial b_R'} = -\frac{R + b_R}{K_1^3}\left(1 + b_R'\right)\left(z' + b_z'\right) + b_z \frac{\partial t_1}{\partial b_R'}\left[2Bb_z \frac{R_1}{PK_1} + \frac{\mu}{2}\left(b - l_c\right)^2\right]$$
$$+ 2Bb_z^2 \frac{t_1}{PK_1}\left(t_2 - R_1 \frac{1 + b_R'}{K_1^2}\right) - Bb_z^2 \frac{R_1}{PK_1^3}\left(z' + b_z'\right)\left(2t_2 - 3R_1 \frac{1 + b_R'}{K_1^2}\right)$$

$$\frac{\partial^2 \mathrm{L}}{\partial z' \partial b'_z} = \frac{R + b_R}{K_1}\left[1 - \frac{\left(z' + b'_z\right)^2}{K_1^2}\right] + b_z \frac{\partial t_3}{\partial z'}\left[2Bb_z \frac{R_1}{PK_1} + \frac{\mu}{2}\left(b - l_c\right)^2\right]$$
$$+ 2Bb_z^2 \frac{t_3}{PK_1}\left(t_1 - R_1 \frac{z' + b'_z}{K_1^2}\right) - Bb_z^2 \frac{R_1}{PK_1^3}\left[2\left(z' + b'_z\right)t_1 + R_1 - 3R_1 \frac{\left(z' + b'_z\right)^2}{K_1^2}\right]$$

$$\frac{\partial^2 \mathrm{L}}{\partial z' \partial R} = \frac{z' + b'}{K_1} + b_z \frac{\mu}{2}\left(b - l_c\right)^2 \left(z' I_2 + b'_z I_3\right)$$
$$+ \frac{Bb_z^2}{PK_1}\left[\left(2t_1 - R_1 \frac{z' + b'_z}{K_1^2}\right)\left(I - \frac{R_1}{P}\right) + 2R_1\left(z' I_2 + b'_z I_3\right) - R_1 I \frac{z' + b'_z}{K_1^2}\right]$$

$$\frac{\partial^2 \mathrm{L}}{\partial z' \partial b_R} = \frac{z' + b'}{K_1} + \frac{Bb_z^2}{PK_1}\left\{\left(2t_1 - R_1 \frac{z' + b'_z}{K_1^2}\right)\left(I_1 - \frac{R_1}{P}\Lambda\right) + R_1\left[2\left(z' I_3 + b'_z I_4\right) - \frac{z' + b'_z}{K_1^2} I_1\right]\right\}$$
$$+ b_z\left(z' I_3 + b'_z I_4\right)\frac{\mu}{2}\left(b - l_c\right)^2 + t_1 \mu\left(b - l_c\right)\frac{b_R b_z}{b}$$

$$\frac{\partial^2 \mathrm{L}}{\partial z' \partial b_z} = Bb_z \frac{R_1}{PK_1}\left(2t_1 - R_1 \frac{z' + b'_z}{K_1^2}\right)\left(2 + b_z \frac{\Lambda_1}{P}\right) + t_1 \frac{\mu}{2}\left(b - l_c\right)^2 + t_1 \mu\left(b - l_c\right)\frac{b_z^2}{b}$$

$$\frac{\partial^2 \mathrm{L}}{\partial z'^2} = \frac{R + b_R}{K_1}\left[1 - \frac{\left(z' + b'_z\right)^2}{K_1^2}\right] + \frac{Bb_z^2}{PK_1}\left\{\left(t_1 - R_1 \frac{z' + b'_z}{K_1^2}\right)\left(2t_1 - R_1 \frac{z' + b'_z}{K_1^2}\right)\right.$$
$$\left. + R_1\left[2\frac{\partial t_1}{\partial z'} - t_1 \frac{z' + b'_z}{K_1^2} - \frac{R_1}{K_1^2} + 2R_1 \frac{\left(z' + b'_z\right)^2}{K_1^4}\right]\right\} + b_z \frac{\mu}{2}\left(b - l_c\right)^2 \frac{\partial t_1}{\partial z'}.$$

The volume constancy condition requires derivatives $\frac{\partial v}{\partial \eta}$ and also various second partial derivatives of v. Here they are:

$$\frac{\partial v}{\partial z} = 0 \quad \frac{\partial v}{\partial b_R} = b_z I_1 \quad \frac{\partial v}{\partial b_z} = RI + b_R I_1 \quad \frac{\partial v}{\partial R} = b_z I$$

$$\frac{\partial v}{\partial z'} = b_z t_1 \quad \frac{\partial v}{\partial b'_R} = b_z t_2 \quad \frac{\partial v}{\partial b'_z} = b_z t_3$$

$$\frac{\partial^2 v}{\partial R \partial z'} = b_z\left(z' I_2 + b'_z I_3\right) \quad \frac{\partial^2 v}{\partial R \partial b'_R} = b_z\left(I_3 + b'_R I_4\right) \quad \frac{\partial^2 v}{\partial R \partial b'_z} = b_z\left(z' I_3 + b'_z I_4\right)$$

$$\frac{\partial^2 v}{\partial b_R \partial z'} = b_z\left(z' I_3 + b_z' I_4\right) \quad \frac{\partial^2 v}{\partial b_R \partial b_R'} = b_z\left(I_4 + b_R' I_5\right) \quad \frac{\partial^2 v}{\partial b_R \partial b_z'} = b_z\left(z' I_4 + b_z' I_5\right)$$

$$\frac{\partial^2 v}{\partial b_z \partial z'} = t_1 \quad \frac{\partial^2 v}{\partial b_z \partial b_R'} = t_2 \quad \frac{\partial^2 v}{\partial b_z \partial b_z'} = t_3$$

$$\frac{\partial^2 v}{\partial z'^2} = b_z \frac{\partial t_1}{\partial z'}$$
$$= b_z\left[RI_2 + b_R I_3 - Rz'^2 J_2 - z'\left(2Rb_z' + b_R z'\right) J_3 - b_z'\left(Rb_z' + 2b_R z'\right) J_4 - b_R b_z'^2 J_5\right]$$

$$\frac{\partial^2 v}{\partial z' \partial b_R'} = b_z \frac{\partial t_2}{\partial z'} = b_z \frac{\partial t_1}{\partial b_R'}$$
$$= b_z\left[-Rz' J_3 - \left(Rb_z' + Rb_R' z' + b_R z'\right) J_4 - \left(Rb_R' b_z' + b_R b_z' + b_R b_R' z'\right) J_5 - b_R b_R' b_z' J_6\right]$$

$$\frac{\partial^2 v}{\partial z' \partial b_z'} = b_z \frac{\partial t_3}{\partial z'} = b_z \frac{\partial t_1}{\partial b_z'}$$
$$= b_z\left[RI_3 + b_R I_4 - Rz'^2 J_3 - z'\left(2Rb_z' + b_R z'\right) J_4 - b_z'\left(Rb_z' + 2b_R z'\right) J_5 - b_R b_z'^2 J_6\right]$$

$$\frac{\partial^2 v}{\partial b_R' \partial z'} = \frac{\partial^2 v}{\partial z' \partial b_R'}, \quad \textit{see above}$$

$$\frac{\partial^2 v}{\partial b_R'^2} = b_z \frac{\partial t_2}{\partial b_R'} = b_z\left[RI_4 + b_R I_5 - RJ_4 - \left(2Rb_R' + b_R\right) J_5 - b_R'\left(Rb_R' + 2b_R\right) J_6 - b_R b_R'^2 J_7\right]$$

$$\frac{\partial^2 v}{\partial b_R' \partial b_z'} = b_z \frac{\partial t_3}{\partial b_R'} = b_z \frac{\partial t_2}{\partial b_z'}$$
$$= b_z\left[-Rz' J_4 - \left(Rb_z' + Rb_R' z' + b_R z'\right) J_5 - \left(Rb_R' b_z' + b_R b_z' + b_R b_R' z'\right) J_6 - b_R b_R' b_z' J_7\right]$$

$$\frac{\partial^2 v}{\partial b_z' \partial z'} = \frac{\partial^2 v}{\partial z' \partial b_z'}, \quad \textit{see above}$$

$$\frac{\partial^2 v}{\partial b_z' \partial b_R'} = \frac{\partial^2 v}{\partial b_R' \partial b_z'}, \quad \textit{see above}$$

$$\frac{\partial^2 v}{\partial b_z'^2} = b_z \frac{\partial t_3}{\partial b_z'}$$
$$= b_z\left[RI_4 + b_R I_5 - Rz'^2 J_4 - z'\left(2Rb_z' + b_R z'\right) J_5 - b_z'\left(Rb_z' + 2b_R z'\right) J_6 - b_R b_z'^2 J_7\right]$$

Auxiliary relationships:

$n \geq 2$

$$\frac{\partial I_n}{\partial z'} = -z' J_n - b'_z J_{n+1}$$

$$\frac{\partial I_n}{\partial b'_R} = -J_{n+1} - b'_R J_{n+2} \qquad \frac{\partial I_n}{\partial b'_z} = -z' J_{n+1} - b'_z J_{n+2}$$

$$\frac{\partial I}{\partial b'_R} = \frac{\partial t_2}{\partial R} = I_3 + b'_R I_4 \qquad \frac{\partial I}{\partial b'_z} = \frac{\partial t_3}{\partial R} = z' I_3 + b'_z I_4$$

$$\frac{\partial I_1}{\partial b'_R} = \frac{\partial t_2}{\partial b_R} = I_4 + b'_R I_5 \qquad \frac{\partial I_1}{\partial b'_z} = \frac{\partial t_3}{\partial b_R} = z' I_4 + b'_z I_5$$

$$\frac{\partial R_1}{\partial z'} = t_1 \quad \frac{\partial R_1}{\partial b'_R} = t_2 \quad \frac{\partial R_1}{\partial b'_z} = t_3 \qquad R_1 = \frac{v}{b_z}$$

$$\frac{\partial t_3}{\partial b'_R} = \frac{\partial t_2}{\partial b'_z} \qquad \frac{\partial t_3}{\partial z'} = \frac{\partial t_1}{\partial b'_z},$$

etc. So that, e.g.,

$$\frac{\partial t_1}{\partial b'_R} = \frac{\partial t_2}{\partial z'} = R\frac{\partial I_3}{\partial z'} + \left(Rb'_R + b_R\right)\frac{\partial I_4}{\partial z'} + b_R b'_R \frac{\partial I_5}{\partial z'}$$
$$= -Rz' J_3 - \left(Rb'_z + Rb'_R z' + b_R z'\right) J_4 - \left(Rb'_R b'_z + b_R b'_z + b_R b'_R z'\right) J_5 - b_R b'_R b'_z J_6$$

$$\frac{\partial t_2}{\partial b'_R} = RI_4 + b_R I_5 - RJ_4 - \left(2Rb'_R + b_R\right) J_5 - \left(Rb'^2_R + 2b_R b'_R\right) J_6 - b_R b'^2_R J_7$$

$$\frac{\partial t_3}{\partial b'_R} = \frac{\partial t_2}{\partial b'_z} = -Rz' J_4 - \left(Rb'_z + Rb'_R z' + b_R z'\right) J_5 - \left(Rb'_R b'_z + b_R b'_z + b_R b'_R z'\right) J_6 - b_R b'_R b'_z J_7$$

$$\frac{\partial t_3}{\partial b'_z} = RI_4 + b_R I_5 - Rz'^2 J_4 - \left(2Rb'_z z' + b_R z'^2\right) J_5 - \left(Rb'^2_z + 2b_R b'_z z'\right) J_6 - b_R b'^2_z J_7$$

$$\frac{\partial t_1}{\partial z'} = RI_2 + b_R I_3 - Rz'^2 J_2 - \left(2Rb'_z + b_R z'\right) z' J_3 - \left(Rb'_z + 2b_R z'\right) b'_z J_4 - b_R b'^2_z J_5$$

so forth.

References

1. J.N. Israelachvili, D.J. Mitchell, B.W. Ninham Theory of self-assembly of hydrocarbon amphiphiles into micelles and bilayers. *J. Chem. Soc. Faraday Trans.* **72**, 1525-1568 (1976).
2. J.N. Israelachvili. *Intermolecular and Surface Forces*, 2nd ed. Academic Press, 1992.
3. J.N. Israelachvili, I. Ladyzinski. The physico-chemical basis of self-assembling structures. In: *Forces, Growth and Form in Soft Condensed Matter: At the Interface between Physics and Biology* (A.T.Skjeltorp, A.V. Belushkin, eds.). Kluwer Academic, 2004, pp.1-28.
4. S. May. A molecular model for the line tension of lipid membranes. *Eur. Phys. J. E* **3**, 37-44 (2000).

EOARD Partner Project 211 – Deliverable 4

1. Introduction

In this last Deliverable we present some results of calculations within the framework of our modification of the opposing forces model (OFM), in particular, with allowance of tilt rigidity of lipid layers. The changes of the membrane shape in the course of peptide binding and further insertion into the membrane are described. The calculation program allows one to investigate the influence of ionic strength, peptide charge and shape, membrane parameters, etc, upon the binding energy and the system energy profile. At the same time, close attention is focused to the revealed limitations and inconsistencies of some points of the OFM that impede its full-value application to the analysis of the translocation process. Thus, apart from necessity of involving the tilt rigidity, the topological and energetic drawbacks of the OFM are shown, as well as its criticality to instabilities of lipid aggregates described within its framework (first of all, of a planar (bi)layer) and to non-locality of some of the OFM basic relationships. The ways of eliminating these drawbacks are pointed out. Proceeding from all this, several promising leads of further research are analyzed in detail.

2. Computational scheme and results

It has become evident in the course of modeling that the extended OFM has several serious drawbacks. In particular, with the typically exploited free energy functional [1] it is unable to produce the unperturbed shape of a planar membrane (see below). Apart from solving these problems theoretically, we managed to achieve the convergence of computational scheme by introducing the tilt rigidity term of ~10 kT into the optimized functional. The volume of the hydrophobic part of the membrane was maintained fixed by applying a standard scheme with Lagrange multiplier l which value was guessed by tries-and-errors. Luckily, the value of l appears to be independent of the electrostatic shielding constant and membrane dimensions, that made the calculations feasible without re-evaluation of l, what is extremely costly computationally.

It is necessary to note that the shape of the membrane in our computations is presented as a series of few cosines; therefore, possible deformations of the membrane are limited to those described as a finite sum of cosines. Particularly, point defects of the lipid packing that require a very large number of Fourier components cannot emerge in calculations. This is a possible reason of the fact that the computational scheme is stable while the underlying energy functional has no stable steady-states in terms of Lyapunov stability analysis.

2.1. Computational details

The program that implements the algorithm of shape variations (see Deliverable 2,3) was used with several modifications. The program searches for five unknown functions, *z*, *b*, *q*, q and f in the interval from 0 to *L*. Each function is expanded in a series of *M* sines or cosines as follows:

$$z(R) = \sum_{i=1}^{M} \tilde{z}_i \cos(w_i R)$$

$$q(R) = \sum_{i=1}^{M} \tilde{q}_i \sin(\nu_i R)$$

$$w_i = \frac{p}{2L} + \frac{p(i-1)}{L}$$

$$\nu_i = \frac{p i}{L}$$

where *M* is the number of harmonics.

The same expansion as for *z* is used for *b* and *q* and the same expansion as for q is used for f. This choice of expansion coefficients allows us to satisfy the desired boundary conditions

$$\left.\frac{\partial q}{\partial R}\right|_{R=0,L} = 0$$

$$z(L) = 0$$

automatically. The same conditions are also applied to other functions. This ensures that that the studied segment of the membrane is smoothly connected to the unperturbed planar bilayer at point *L* and there is no tilt of lipid directors at point *R*=0 (prohibited by symmetry). The derivatives of all functions are computed analytically. The program performs the minimization in the space of expansion coefficients $\tilde{z}_i$, $\tilde{b}_i$, $\tilde{q}_i$, $\tilde{q}_i$ and $\tilde{f}_i$ using the method of simple unconditional minimization (see Deliverables 2,3).

The electrostatic interaction between the charged peptide and the membrane is taken in the form

$$E_{elect} = \sum_{j=1}^{N_p} \int_0^L \frac{1}{4pe_0} \frac{q_e s_u(R)}{d_j(R) a_h(R)} \exp\left(\frac{d_j(R)}{d_0}\right) dR ,$$

where N_p is the number of charges of the peptide, e_0 is the vacuum dielectric constant, q_e is the elementary charge, $s_u(R)dR$ is the infinitesimal area of the interface at the level of the polar head groups, a_h is the area per lipid at the same level, d_0 is the shielding constant for electrostatic interactions (similar but not identical to the Debay length in continuous electrolyte, and d_j is the distance from the j-th charge of the peptide to a given element of the membrane surface:

$$d_j(R) = \sqrt{X_{\text{int}}(R)^2 + \left(Z_{\text{int}}(R) - z_j\right)^2},$$

where X_{int} and Z_{int} are the coordinates of heads of the lipids anchored at point R; z_j is the coordinate of the j-th charge of the peptide (the latter is positioned at x=0). Multiple charges are positioned evenly along the peptide as described below.

The close contact of the peptide with the membrane is modeled by the "nearly hard core" potential assigned to the peptide:

$$E_{VDW}(r) = \begin{cases} \dfrac{k}{(r - r_{core})^{12}}, & r > r_{core} + d \\ K, & r \le r_{core} + d \end{cases}$$

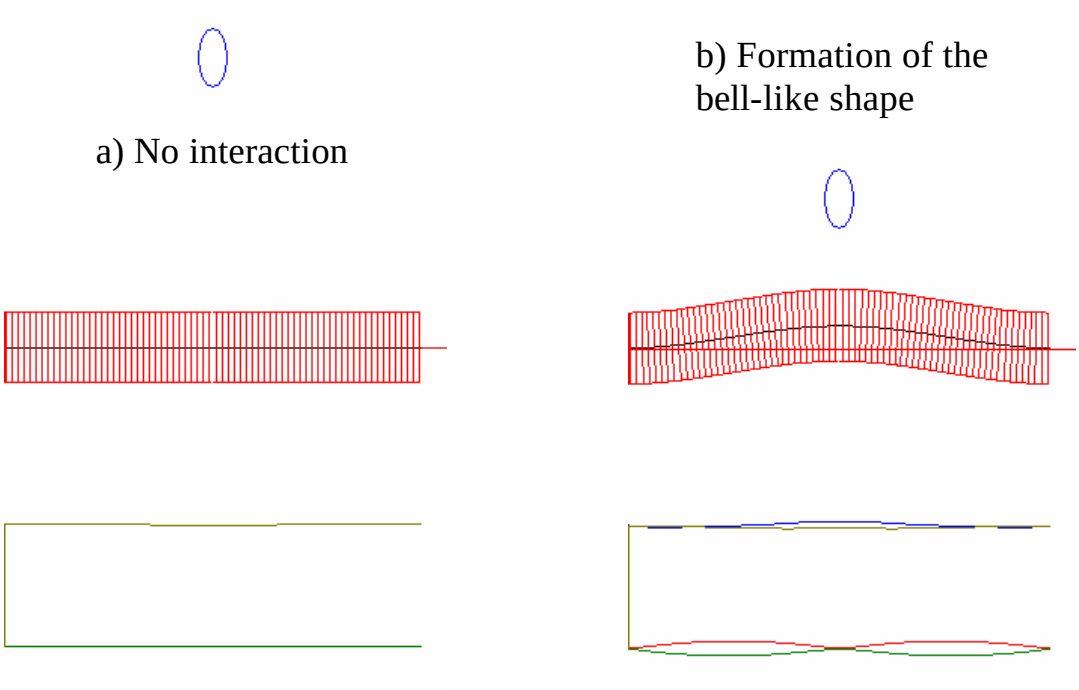
where r_{core} is the radius of the hard sphere that surrounds the lowermost charge of the peptide, k and K are adjustable parameters, r is the distance between the given discrete portion of the interface and the lowermost charge of the peptide, d is a very small shift. In typical simulations k=1, K=10^{10}, $d = 10^{-8}$. The common problem of modeling the constraints in variation methods is the "dead lock" which often occurs after the first contact with the hard core. Our form of interaction produces a behavior which is almost indistinguishable from that under the true hard core impermeability criterion and allows us to overcome computational complications.

2.2. Shape of the membrane

The following series of snapshots illustrates the principal behavior of the membrane with the peptide approaching. The value of the electrostatic shielding constant is kept very high (25 Å) in order to make the shape variations more pronounced. At the first stage the membrane remains unaffected in its equilibrium state (Fig. 1a). When the peptide goes closer to the charged interface the electrostatic attraction begins to curve the membrane toward the peptide forming a bell-like shape maintained by the balance of electrostatic and elastic forces (Fig. 1b). Eventually the

membrane collides with the hard core of the peptide (Fig. 1c). Starting from this stage the forces of hard-core van-der-Waals (VDW) repulsion and electrostatic attraction act in opposite directions. As a result the peptide “digs” a central well in the upward bell-like shape of the membrane forming the "M-like" shape (Fig. 1d). Since the peptide goes close to the membrane baseline, the average curvature of the M-shaped membrane decreases and the associated elastic energy becomes small in comparison to early bell-like structures. However, the electrostatic interaction energy between the peptide and the charges membrane interface remains high. As a result, the M-like structures correspond to the minimal energy of the membrane-peptide system and are likely to be observed in reality. Further approach of the peptide forces the membrane to follow it. The wings of the M-shape become flattened and the central depression under the peptide becomes the dominant feature (Fig. 1e). Further movements of the peptide lead to the deepening of this depression and to dramatic increase of the elastic deformation energy (Fig. 1f). The plots under the schemes of the membrane show the deviations of the lipid director length (blue – the upper monolayer, grey – the lower) and the director tilt angles (red – the upper monolayer, green – the lower). The deviations are not scaled and serve for comparative purposes only.

The value of the shielding constant d=25 Å is not realistic since it corresponds to the environment of very low ionic strength. More realistic values are d=3÷6 Å. For such values the overall picture of the membrane deformations remains the same but becomes much less pronounced.



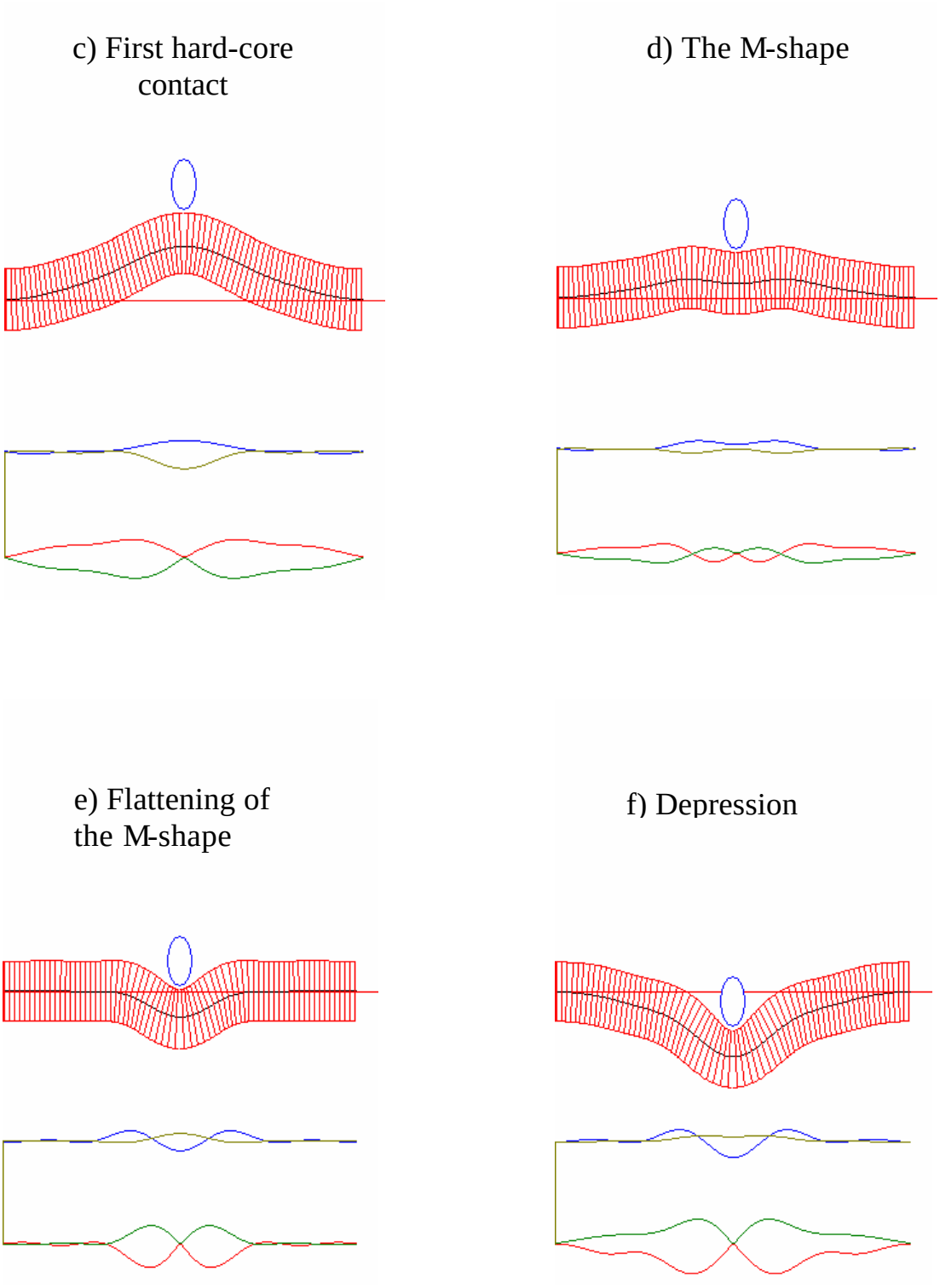


Fig. 1. Snapshots of the membrane shape for various positions of the peptide (see the text). Radius of the modeled part of the membrane is 140 Å (shown cross-section is 280 Å wide). The peptide is modeled by a hard 10 Å sphere with the single point charge in its center. Vertical scale is double-stretched to make the shape variations more pronounced.

2.3. The influence of ionic strength

We calculated the potential energy profiles of binding for different values of the shielding constant. The results are shown in Fig.2. It is clearly seen that the shielding constant influences the shape and the depth of the energy profile dramatically. Decrease of the shielding constant decreases the binding energy in non-linear fashion (Fig.2). The most physiologically relevant value of the shielding constant is rather hard to estimate since it includes not only the conventional Debay factor, but also the effective influence of the polarization effects and local conformational changes in the membrane and the peptide. The value of 5Å can be assigned as a reasonable estimate that leads to the binding energy of 7.7 *kT*. This value is quite close to experimentally determined binding energy per peptide. It is necessary to emphasize, however, that this value is obtained for a single point charge approaching to the membrane, thus it can be considered as a rough estimate only.

2.4. The influence of the peptide charge

To study the influence of distributed peptide charge on the binding energy, we adopted the following model of the peptide. The latter is represented by a vertical line of length L=15 Å. The first charge is located at the lower end of the line, while others are distributed at even distances along the line. The hard-core repulsion was calculated for the lowest charge only since other charges do not contact with the membrane surface directly. The shielding constant is fixed at the value of 5Å. It can be seen in Fig.3 that the increase in the peptide charge leads to gradual (almost linear) increase in the binding energy.

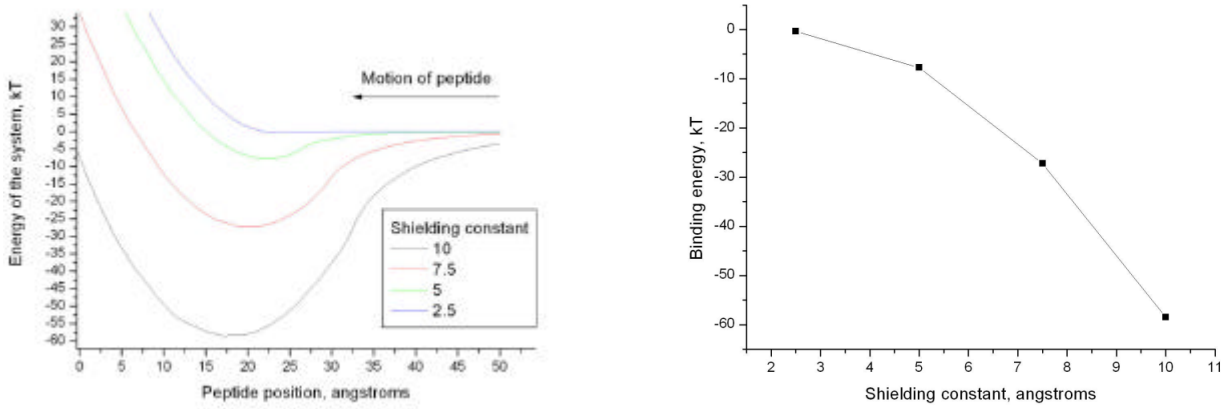


Fig. 2. Energy profiles and binding energy at different levels of shielding.

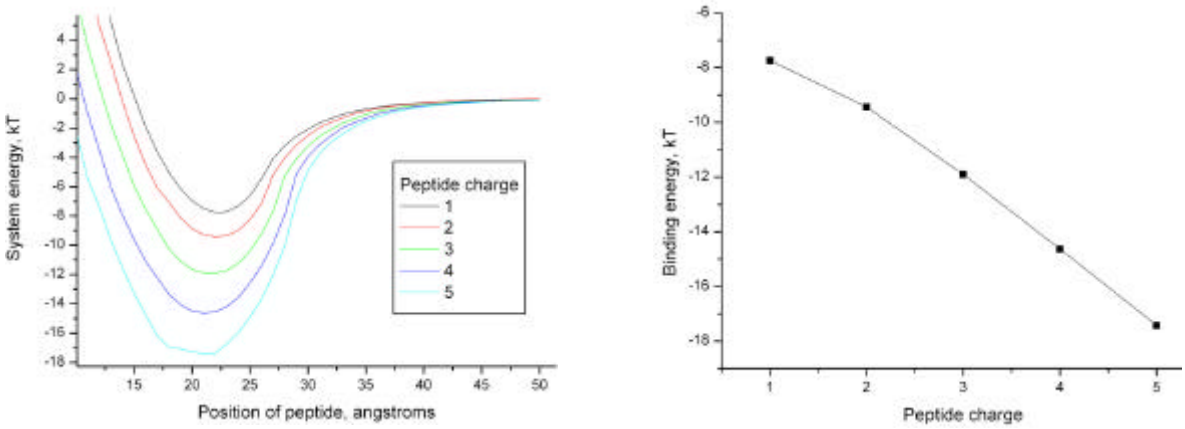


Fig. 3. Energy profiles and the binding energy for various peptide charges.

3. The OFM model should be further modified to describe pore formation

The ultimate goal of the current study is to simulate the formation of the pore in the membrane caused by the influence of the CPP. The most probable mechanism of pore formation is formation of the semi-toroidal micelle-like rim around the water-filled pore. This excludes unfavorable contacts of lipid tails with water and forms a hydrophilic channel for the permeating peptide. Formation of such a membrane shape implies that the packing of the lipid tails is not uniform and differ considerably in the micellar rim and in the distant parts of planar bilayer. It was claimed in several publications (see Refs. in Deliverables 2,3) that the modified semi-phenomenological OFM model is applicable to membrane shapes with non-uniform chain packing. Particularly, the energy of the semi-micellar bilayer edge was computed [1]. However, there is one extremely important assumption in these works, not emphasized by the authors but critical for applications of the modified OFM model. In all mentioned works the existence of a stable symmetric bilayer is *postulated*, not *derived* from the first principles. This approach looks sufficient if the geometry of the membrane system under study is simple enough and known from experiment. However, it fails if there is no *a priori* knowledge of the membrane shape. The binding problem implies that the membrane remains in the bilayer form but bends to accommodate the peptide. This process can be classified as manageable by the OFM-like model we use. However, the problem of pore formation is an unfavorable case when there is no notion of the shape of intermediate states of the membrane.

Let us examine the problem in more details. The transition from a planar bilayer to a semi-toroidal pore changes the topology of the membrane surface qualitatively. In rigirous topological terms, a planar membrane is isomorphous to a sphere, while a pore is isomorphous to a toroid. Within the OFM the membrane is treated in continuous approximation and the lipid

directors are "anchored" on the membrane midline. However, the latter becomes discontinuous when the membrane evolves from a bilayer to a pore. Since all energy contributions are expressed in terms of the midline, they are not topologically invariant. The simplest illustration is given in Fig. 4a. In the "ideal" membrane edge capped by the micellar region the directors of all lipids that form the micellar cap originate from the same point (the origin of the membrane midline). At the same time, in the bilayer part, only one lipid per monolayer can originate from the same point. In Ref.1 this problem was solved by *assuming* that the micellar cap is already formed. Then the system was split into two topologically different parts –the bilayer and the micelle and *different* energy functionals were written for each of them. However, we *can not* go this way because we have to describe the *formation* of the micellar part, that is, a smooth transition from the geometry of a bilayer to the geometry of a micelle. The only possible way to do this is illustrated in Fig. 4. If the number of discrete points that represent the membrane is large enough, then the difference between these two representations becomes negligible.

However, it is still impossible to describe a smooth transition from the bilayer to the pore. In order to form the "seed" of the pore, the origin of the midline (bold dot in Fig.4) should move away from zero (symmetry axis of the future pore). In this case there is no lipid "anchored" between the pore axis and the midline origin. From the physical point of view this means that the tails of the lipids closest to the axis will repack themselves to fill the gap. Since this packing will obviously be not ideal, the system energy will increase. However, appearance of such a gap becomes fatal for the OFM. The energy of the membrane is computed as an integral over the midline, and the regions where the latter is absent contribute nothing to the total energy. In other words, the "missed volume" shown in Fig.4 has zero energy in terms of the OFM. The system is absolutely invariant to the size and shape of this region, what contradicts physical reality.

The next step of pore formation is the rapture of the interfaces. This allows water to fill the pore and to contact with the hydrophobic tails of the lipids. In reality, such a contact is extremely unfavorable and the first lipids of the upper and lower interfaces will approach each other to eliminate the hydrophobic mismatch. However, the OFM contains *no explicit terms* describing the hydrophobic mismatch – the system does not "feel" that the tails are exposed to water.

There is one more energy term missed in the OFM. The heads of the lipids are often charged and repel each other. This interaction within *single* interface is implicitly described by the B/a_h term of the OFM. However, if two *separate* interfaces are located close to each other, there is a strong electrostatic interaction between them. This interaction should be described explicitly but it is missed in the OFM. Introduction of this additional energy term is far from trivial. The micellar cap formally consists of two parts – one from the upper and another from the lower interface. An explicit electrostatic interaction should act between these parts, but at the

same time the interface of the micelle is topologically continuous and should be described by the implicit term itself.

All the three energy components missed in the OFM are schematically shown in Fig.4.

4. Internal inconsistencies of the extended OFM

4.1. The tilt energy

It is well known that lipid bilayers possess some rigidity to tilt deformations that are often described in the linear approximation by the tilt modulus (see e.g. [2]). The physical meaning of tilt rigidity in terms of the OFM originates from the fact that the fluctuating hydrophobic tails of the lipids cannot cross the interface with water in order not to be exposed. If the director of a lipid is tilted from the normal of the hydrophobic-water interface, a certain part of the space available for the fluctuating hydrocarbon chain in normal orientation becomes excluded. This leads to some loss of conformational freedom and, as a result, to the change in the entropic contribution to the free energy of the lipid. The energy term associated with this effect can be written approximately as $E_{tilt} = k_{tilt} a^2$, where a is the angle between the normal to the hydrophobic-water interface and the lipid director, and k_{tilt} is the tilt modulus. It was shown that the tilt rigidity of the bilayers is quite high and the tilt modulus reaches ~10 kT [2]. However, this term was not included into the modified OFM. We have found that this term is vital for maintaining the stability of the membrane. If this term is omitted, certain "unfortunate" situations can appear during the optimization when the lipids orient themselves almost parallel to the interface and get stuck in this geometry (see Fig.5). This configuration is clearly nonphysical, since the hydrocarbon tails have significant thickness and will be exposed to water if a is close to 90°. We introduced the tilt term into our simulations and assigned the value 10 kT to the tilt modulus.

4.2. The challenge of non-locality

In order to find the area and the volume per lipid the following expressions are used (for the sake of clarity we here consider a quasi-one-dimensional case that does not restrict the problem generality):

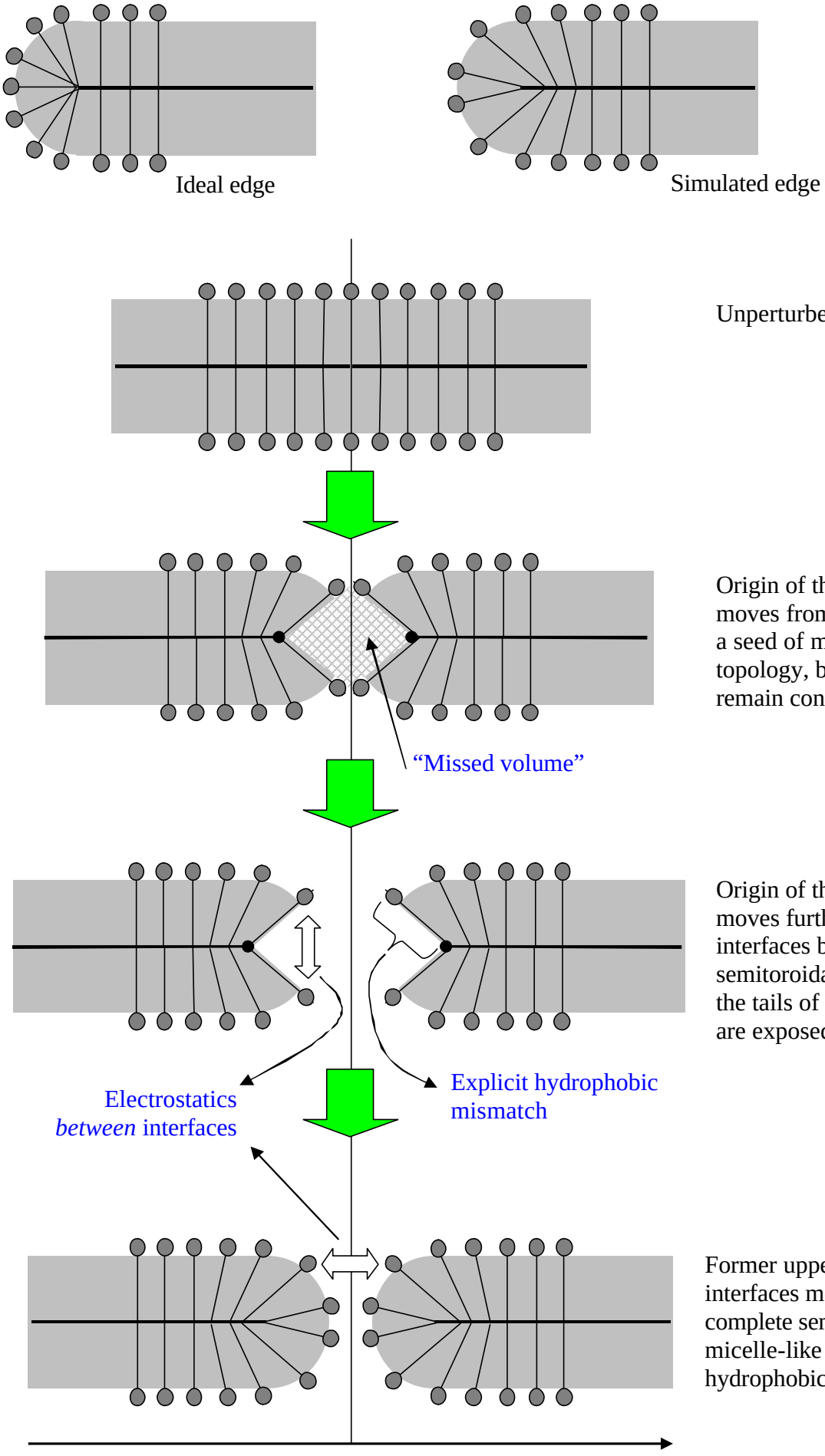

Ideal edge
Simulated edge
Unperturbed bilayer
Origin of the midline moves from zero, forming a seed of micellar topology, but the interfaces remain continuous.
"Missed volume"
Origin of the midline moves further, the interfaces break, the semitoroidal rim forms, but the tails of lipids at the edge are exposed to water.
Electrostatics *between* interfaces
Explicit hydrophobic mismatch
Former upper and lower interfaces merge forming a complete semi-toroidal micelle-like rim with no hydrophobic mismatch.

Fig. 4.

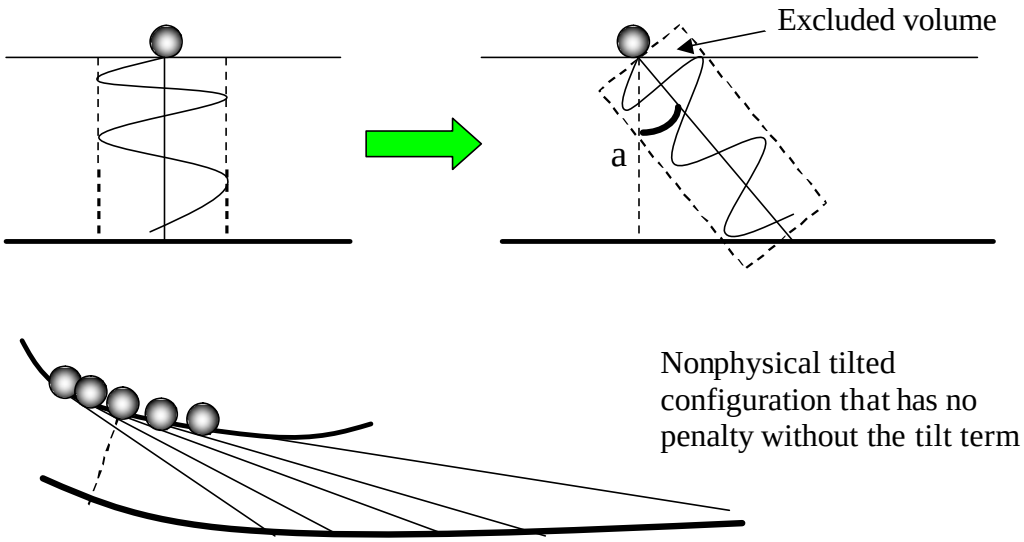


Fig. 5.

$$a_i(R) = \int_R^{R+\Delta R} s(R')dR' \approx s(R)\Delta R$$

$$u = \int_R^{R+\Delta R} v(R')dR' \approx v(R)\Delta R$$

what is approximately valid if $\Delta R << R_{\min}$, where R_{min} is the minimal local curvature radius of the monolayer in the course of membrane deformation. For large deformations this approximation is not valid and, strictly speaking, both equations should be solved to find ΔR. The latter is the radial size of an average lipid anchored at point R of the midline. It is easy to show that for an unperturbed bilayer $\Delta R \sim b_0$, where b_0 is the equilibrium thickness of a monolayer (typically ~ 10 Å). This means that an object that spans the radial dimension of ~10 Å is treated as an infinitesimally small volume which is integrated over R to obtain the total energy of the monolayer F_M (see e.g. Eq.(12) in Deliverable 2).

However, it has turned out to be a rather rough approximation. In fact, in the micellar part of the system $\Delta R \sim R_{\min}$ and the area per lipid should be expressed by the functional $a_i(R) = \int_R^{R+\Delta R(R)} s(R')dR'$. In this case the minimization of the functional F_M can no longer be performed by the standard Euler-Lagrange method. The reason is that the later can only be applied if the functional of form $F(x,t) = \int f(x,\dot{x},t)dt$ is minimized while in our case we have a

"functional of a functional". Minimization of such a complex construct is itself a non-standard and challenging mathematical problem. Additional complication comes from the fact the computational methods of optimization of such complex "double" functional are underdeveloped (our results to these specific issues are now in progress and will be presented elsewhere).

4.3. The ambiguity of the constant volume constraint

A physical formulation of our problem implies the "flat unperturbed membrane" as a reference state. This trivial phrase contains a subtle pitfall, however. A real biological membrane (say, a liposome of large radius) has a fixed total volume of the hydrophobic core, conserved with high accuracy. We consider a much smaller piece of the membrane and imply that it remains unperturbed infinitely far from the point of perturbation by the peptide. In practice, we consider the fraction of the membrane spanning from $-l$ to l, where l~100-200 Å. In the unperturbed state this piece contains N lipids and has volume V of the hydrophobic core. Perturbation of the membrane causes the changes in chain packing and redistribution of lipids along the membrane. Suppose that the membrane takes a bell-like shape (see Fig.6). The lipids located at points l and $-l$ tend to move toward the center of perturbation. Several kinds of the boundary conditions are then possible:

(i) Impermeable walls – no lipids can cross the boundaries, $l = const$, $N = const$, $V = const$. If the membrane is deformed, its area increases and substantial surface tension appears. The energy of this artificial strain is added to the "intrinsic" deformation energy of the membrane introducing some systematic error.

(ii) "Transparent" walls – the lipids can freely cross the boundaries, $l = const$ but N and V are not conserved. This seems to be the most physically correct version – the lipids can leave the given volume but accumulate outside in some other parts of a huge liposome. In practice, this leads to the fact that *all* lipids tend to leave the volume. The reason of this is that the energies in the OFM are always positive and, therefore, the state with no lipid is, erroneously, the global energy minimum. Thus, this type of boundary conditions is not applicable.

(iii) Moving impermeable walls – no lipids can cross the boundaries, $N = const$, $V = const$, but the value of l can be adjusted to reduce the tension caused by deformation. This type of boundary conditions is the most preferable since it retains the integrity of the membrane and eliminates artificial strain. Unfortunately, moving boundaries are quite hard to implement computationally because the discrete lattice used in simulations should be redefined at each iteration. This results in numerous specific problems and greatly increases computational burden. That is why we decided not to implement this method until we improve the accuracy of the

membrane model to the point when the error introduced by additional mechanical strain will become detectable.

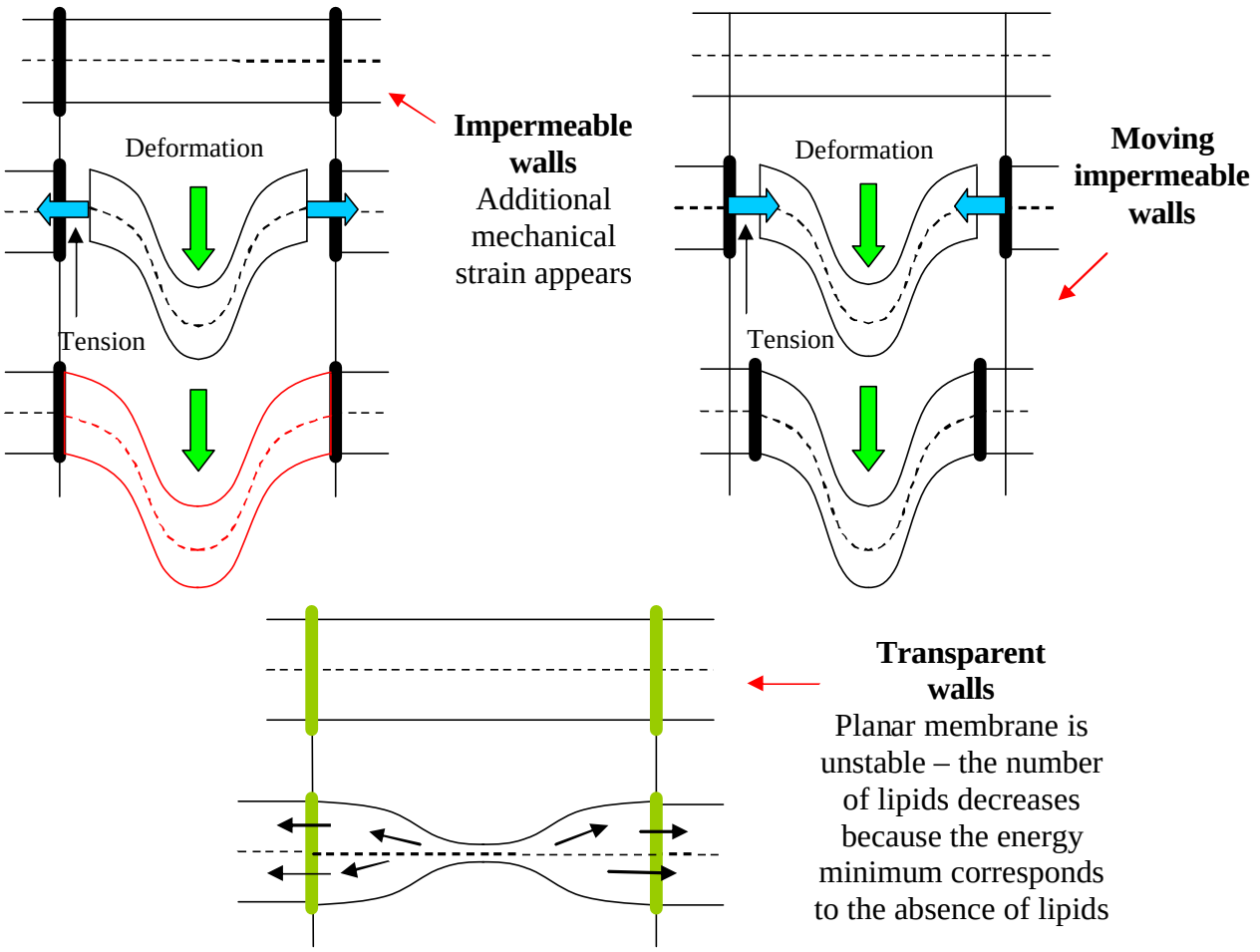


Fig. 6.

4.4. Stability problem

In this subsection we elucidate a crucial point at issue for the OFM validity, namely, the stability of lipid aggregates within the OFM framework.

To illustrate the problems arising in the procedure of minimization of the free energy functional, we first consider a Lagrangian $\mathbf{L}(y(x), y'(x))$ entering functional $\int \mathbf{L}(y, y')dx$ and corresponding Euler-Lagrange equations $\frac{\partial \mathbf{L}}{\partial y} = \frac{d}{dx}\frac{\partial \mathbf{L}}{\partial y'}$, or

$$y'' = \left(\frac{\partial^2 \mathbf{L}}{\partial y'^2}\right)^{-1}\left(\frac{\partial \mathbf{L}}{\partial y} - y'\frac{\partial^2 \mathbf{L}}{\partial y \partial y'}\right) \qquad (1)$$

provided that $\mathbf{L}$ does not depend on x explicitly. Obviously, the "stationary" solution y_s of Eq.(1) is defined by the equation

$$\left.\frac{\partial \mathbf{L}}{\partial y}\right|_s = 0 . \qquad (2)$$

Here and below the subscript s labels the functions taken at $y = y_s, y' = 0$, that is, for $f(y, y')$ $f_s \equiv f(y_s, 0)$. In the standard linear analysis of stability one re-writes (1) as a set of two first-order equations for variables $y_1 = y,\ y_2 = y'$:

$$\begin{cases} \dfrac{dy_1}{dx} = y_2 \\ \dfrac{dy_2}{dx} = \left(\dfrac{\partial^2 \mathbf{L}}{\partial y_2^2}\right)^{-1}\left(\dfrac{\partial \mathbf{L}}{\partial y_1} - y_2 \dfrac{\partial^2 \mathbf{L}}{\partial y_1 \partial y_2}\right) \equiv F_2(y_1, y_2). \end{cases} \qquad (3)$$

and solves the characteristic equation of set (3) linearized in the vicinity of point $(y_1 = y_s,\ y_2 = 0)$. The corresponding roots read:

$$k_{1,2} = \frac{1}{2}\left.\frac{\partial F_2}{\partial y_2}\right|_s \pm \sqrt{\frac{1}{4}\left(\left.\frac{\partial F_2}{\partial y_2}\right|_s\right)^2 + \left.\frac{\partial F_2}{\partial y_1}\right|_s} ,$$

so that stability $\left(\mathrm{Re}\, k_{1,2} < 0\right)$ requires the following inequalities to hold:

$$\left.\frac{\partial F_2}{\partial y_2}\right|_s < 0, \quad \left.\frac{\partial F_2}{\partial y_1}\right|_s \le 0.$$

However, it is easy to find that

$$\frac{\partial F_2}{\partial y_2} = \left(\frac{\partial^2 \mathbf{L}}{\partial y_2^2}\right)^{-2}\left[-y_2 \frac{\partial^3 \mathbf{L}}{\partial y_1 \partial y_2^2}\frac{\partial^2 \mathbf{L}}{\partial y_2^2} - \frac{\partial^3 \mathbf{L}}{\partial y_2^3}\left(\frac{\partial \mathbf{L}}{\partial y_1} - y_2 \frac{\partial^2 \mathbf{L}}{\partial y_1 \partial y_2}\right)\right]$$

so that $\left.\frac{\partial F_2}{\partial y_2}\right|_s = 0$ because $\left.\frac{\partial \mathbf{L}}{\partial y_1}\right|_s = 0$, see Eq.(2). This means nothing else than stationary states of the variational problem with such kind of Lagrangian are *always unstable* (as it should be expected from a corresponding mechanical analogue). In our context, we thus conclude that the OFM is not able to ensure the existence of a stable planar (bi)layer!

One of the ways out of the situation is to insert an artificial decay term proportional to $(-y')$ in the r.h.s. of (1), exactly as the friction term is introduced in equations of classical mechanics. Of course, in our case physical meaning could hardly be attached to such "friction". It is worth to note, however, that this term affects neither the stationary solution y_s nor the stability domain in the parameter space (in particular, the coefficient $b > 0$ in this formal term $-b\,y'$ can be even infinitesimal). Indeed, if we suppose that now

$$F_2(y_1, y_2) \equiv \left(\frac{\partial^2 \mathbf{L}}{\partial y_2^2}\right)^{-1}\left(\frac{\partial \mathbf{L}}{\partial y_1} - y_2\frac{\partial^2 \mathbf{L}}{\partial y_1 \partial y_2}\right) - b\,y_2$$

in (3), then we always have $\left.\frac{\partial F_2}{\partial y_2}\right|_s = -b < 0$, y_s remains unchanged, and the stability domain is defined by the (independent of b !) condition:

$$\left.\frac{\partial F_2}{\partial y_1}\right|_s \le 0, \quad \text{or} \quad \left.\left[\left(\frac{\partial^2 \mathbf{L}}{\partial y_1^2}\right)\left(\frac{\partial^2 \mathbf{L}}{\partial y_2^2}\right)^{-1}\right]\right|_s \le 0\,. \tag{4}$$

Return now to the simplest case of a lipid layer within the extended OFM. To analyze the problem, there is no need in all the five sought functions of our model of the membrane. We can come to relevant conclusions by considering a quasi-one-dimensional model of a lipid layer in which all the lipids are perpendicular to the baseline, and the latter is fixed straight. Then the only variable under consideration is the lipid length b (so that $y = b, y' = b'$). In the reduced (dimensionless) units introduced before the corresponding functional reads

$$\Phi[b] = \int_0^{l(b)}\left[\sqrt{1+b'^2} + \frac{Bb^2}{\sqrt{1+b'^2}} + m\left(b - l_c\right)^2 b\right]dx + l\int_0^{l(b)} b\,dx$$
$$\equiv \int_0^{l(b)} L(b, b')dx + l\int_0^{l(b)} b\,dx$$

where the second integral represents the lipid volume constancy condition $\int_0^{l(b)} b dx = V$, l is the Lagrange multiplier, $m = (1-B)/(1-l_c)$, and $l(b)$ is the layer length ("movable walls"); here we omit the bars over dimensionless parameters, see Deliverable 2. The Euler-Lagrange equation reads

$$b'' = \left(\frac{\partial^2 L}{\partial b'^2} \right)^{-1} \left(\frac{\partial L}{\partial b} - b' \frac{\partial^2 L}{\partial b \partial b'} + l \right)$$

and its stationary solution is defined by the equation $l + \left.\frac{\partial L}{\partial b}\right|_{b'=0} = 0$, i.e. by

$$-\frac{3}{2} m b_s^2 - 2 b_s (B - m l_c) - \frac{1}{2} m l_c^2 = l \qquad (5)$$

which can be also regarded as an equation to determine l, for in this case b_s can be found from the following simple considerations. Precisely, for the planar layer the condition $\int_0^{l(b)} b dx = V$ immediately results in the relationship $l(b_s) = V / b_s$. Recall now that $\int_0^{l(b)} L dx$ is nothing else than $\int_0^{l(b)} \frac{f_0}{\nu} b dx$. If $b = b_s$, then this integral is reduced simply to $\frac{f_0}{\nu} b_s l(b_s) = \frac{f_0}{\nu} V = N f_0$ where N is the number of lipids. Therefore, for a planar layer the minimum of the functional corresponds to the minimum of f_0, that is $b_s = b_0$; in our dimensionless units $b_0 = 1$ (see Deliverable 2). In particular, from (5) one has that $l = -2B + \frac{1-B}{1-l_c}\left(2l_c - \frac{l_c^2 + 3}{2} \right)$.

The stability condition (4) now reads

$$\left[\left(\frac{\partial^2 L}{\partial b^2} \right) \left(\frac{\partial^2 L}{\partial b'^2} \right)^{-1} \right] \Bigg|_{\substack{b=1 \\ b'=0}} \leq 0.$$

As $L_{bb}(1,0) = \frac{3 - B - 2l_c}{1 - l_c}$ and $L_{b'b'}(1,0) = 1 - B$, we have

$$\frac{3-B-2l_c}{(1-l_c)(1-B)} \le 0. \qquad (6)$$

Fig.7 shows the quantity $(\partial F_2/\partial y)\big|_s$ as a function of parameters B, l_c and the corresponding stability domain defined by (6).

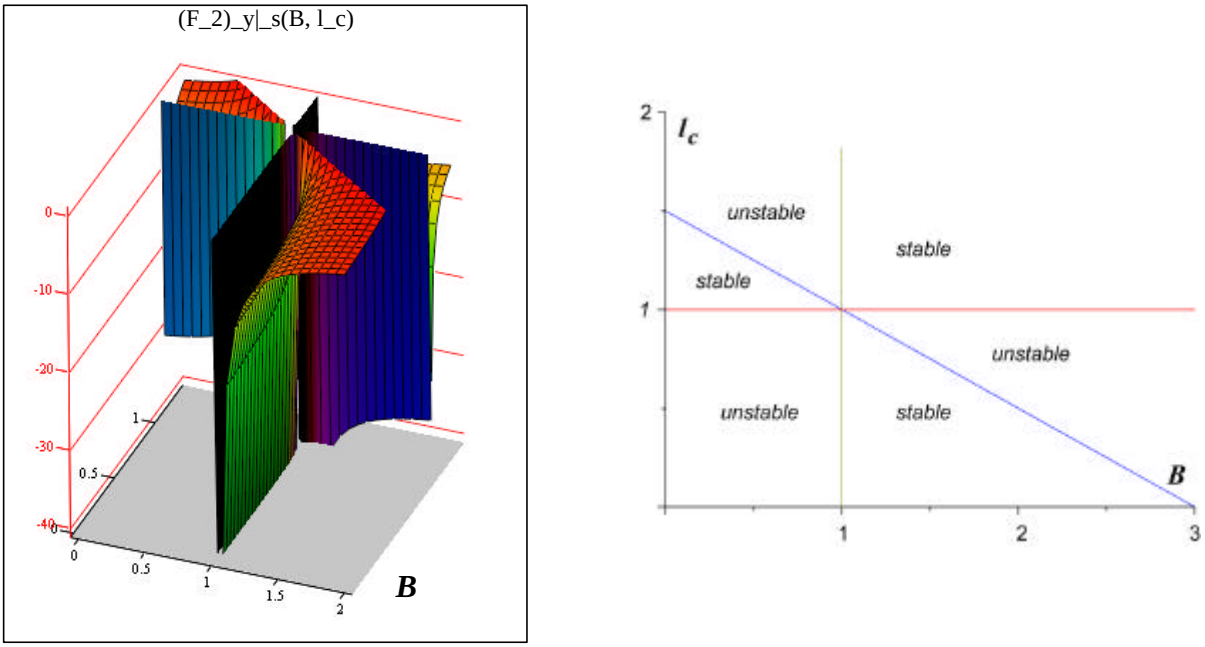


Fig. 7.

The return of the parameters to their initial dimensions is not trivial because b_0 itself is defined as a root of the cubic equation $2t\,vb_0^3 + (B - 2t\,vl_c)b_0^2 - g\,v^2 = 0$ (see Deliverable 2, Eq.(13)); here the coefficients *are bar-less again*. Of course, this conversion can be performed numerically. An example for the parameter values typically used for lipid membranes is shown in Fig.8. One can see how delicate the problem of stability of the planar layer is.

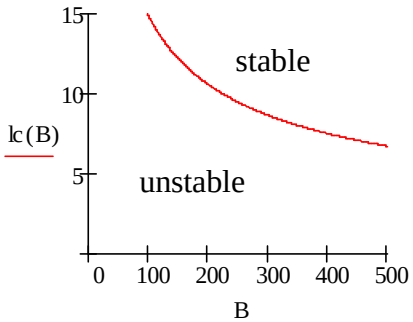


Fig. 8.

$t = 0.1\ kT/\text{Å}^2$; $g = 0.12\ kT/\text{Å}^2$; $v = 432\ \text{Å}^3$.

5. Research perspectives

The present pilot project allowed us to evaluate one of the possible approaches to description of the CPP permeation. This approach is based on the modified semi-phenomenological OFM-like model. This model, to our knowledge, is the most detailed among available continuous membrane models and pretends to describe non-uniform chain packing of various membrane topologies. However, we revealed fundamental obstacles which discouraged us from using this model "as is", without substantial modification or even development of new concepts. We see the following three directions of research that can be further developed proceeding from the results of the project:

A. Development of a continuous semi-phenomenological model, free from the drawbacks of OFM-like models.

Requirements:

1. The energy function contains both surface and volume terms sensitive to the local geometry.
2. The lipid directors are anchored on the hydrophobic-polar surface that remains continuous in all possible membrane perturbations and rearrangements.
3. The volume energy terms allow for possible not ideal packing of lipid tails and for interactions between different monolayers.
4. The relationship between the area and volume of the average lipid anchored in the given point is non-local. It depends on the geometry of the certain membrane segment spanned by the lipid, not only on the geometry at the anchoring point.
5. Additional surface terms describing long-range electrostatic interactions should be added.
6. A flat unperturbed monolayer is a stable steady-state without any additional constraints or assumptions.

Benefits:

1. Simple enough to find the optimal membrane shape of a perturbed membrane by low-cost computational techniques.
2. The variety of parameters that correspond to real membranes and CPPs can be testes.

Limitations:

1. It is not clear if all the requirements can be met simultaneously.
2. The results obtained are only semi-quantitative and depend strongly on the quality of empirical parameters.
3. Hard to correlate some empirical parameters with experimental data.
4. Can never answer the question about atomic-scale details of permeation.

B. Development of a highly simplified discrete molecular-level model of the membrane.

Requirements:

1. Each lipid is modeled as a simple rod-like object.
2. An empirical force field should be developed to describe the interactions between the rods in order to mimic the basic behavior of lipids.
3. The model should mimic only the very basic mechanical properties of the membranes, no need to describe smaller details.
4. The charges of the polar heads are included explicitly as well as electrostatic interaction with the peptide.

Benefits:

1. Eliminates the topological problems of the OFM-like models because the lipids are modeled at the molecular level.
2. Can model *any* shape of the membrane, provided that a correct force field is supplied.
3. Specific interactions of the peptide with the membrane (like strong binding of arginines to phosphates) can be taken into account by additional energy terms.
4. Can be used to model the permeation in dynamics, not just to evaluate the energy profiles of permeation.
5. It is much cheaper computationally than molecular dynamics, allowing longer simulations and experiments with various parameters.

Limitations:

1. The development and validation of the force field is a quite tedious task.
2. The force fields for rod-like particles are inconsistent with existent molecular simulation software. Modification of the existing programs or *de novo* design of the simulation package is required.
3. The results obtained are still semi-quantitative only and depend strongly on the quality of the force-field.

C. Molecular simulations at the atomic or nearly-atomic level.

Requirements:

1. The methodology of such simulations is well known in general but should be adapted to a particular task.
2. The protocol of pulling the peptide through the membrane should be developed because spontaneous permeation is beyond the reachable timescale.
3. The method should be used in conjunction with analytical techniques to analyze the free energy profiles of permeation and to calculate the transition rates.

Benefits:

1. Well-defined methodology and extensively tested force-fields for the membrane and the peptide.
2. Atomic-level details of all binding and permeation events.
3. Explicit comparative tests of real peptides and membranes are possible.
4. We have established contacts with a world-leading molecular dynamics group in Groningen. It is possible to negotiate the usage of their experience and technical capabilities.

Limitations:

1. The method of molecular dynamics is extremely demanding computationally in comparison with previous two ones. Supercomputers/clusters are needed.
2. The time scale of simulations is limited to ~100ns what is insufficient to observe spontaneous permeation events. The "pulling" should be used, allowing to produce the approximate free energy profiles of permeation only.

References

1. S. May. A molecular model for the line tension of lipid membranes. *Eur. Phys. J. E* **3**, 37-44 (2000).
2. S. May, Y. Kozlovsky, A. Ben-Shaul, M.M. Kozlov. Tilt modulus of a lipid monolayer. *Eur. Phys. J. E* **14**, 299-308 (2004).

Summary

The main purpose of the present project was a theoretical investigation of the possibility of translocation of some positively charged peptides (so called cell-penetrating peptides, CPPs) through artificial phospholipid bilayer membranes and evaluation of the influence of different physicochemical factors upon this process.

As a result of analysis of the current, often contradictory experimental data it has become clear that reliable conclusions on the feasibility of the given non-standard phenomenon can be obtained only on the basis of its physical modeling within a microscopic (to reasonable extent) approach. Proceeding from available computational means, we have chosen the known (and widely used in current literature on membrane systems) opposing forces model which is considered as the most suitable for describing large deformations of lipid aggregates and their complexes with proteins.

We have developed an OFM version aimed at a unified and continuous description of the process of binding and translocation of the peptide, including complex changes of the system geometry. In fact, a new approach to the minimization of the free energy functional of systems with essential non-locality and a corresponding computational formalism is constructed. Computer programs are created which allows one to elucidate the shape of a peptide-membrane complex and to calculate the binding energy and energy profile of the system under further insertion of the peptide into the membrane, to investigate the dependences of these characteristics on the ionic strength, peptide charge, membrane parameters, etc.

At the same time, in the course of project execution we have revealed a series of essential drawbacks and inconsistencies of the OFM which hinder its application for a full translocation process, especially at the pore formation stage. We present a detailed analysis of these drawbacks and formulate the principles of construction of a distinctly improved phenomenological model, capable of correct description of membrane rearrangement topology. On this basis we develop the prospective trends of further research, rating them on their success criteria.